北京大学物理学丛书

The Series of Advanced Physics of Peking University

热大爆炸宇宙学

北京大学出版社 PEKING UNIVERSITY PRESS

北京大学物理学丛书

热大爆炸宇宙学

俞允强 著

北京大学出版社
北 京

图书在版编目(CIP)数据

热大爆炸宇宙学/俞允强著. —北京:北京大学出版社,2001.4
ISBN 7-301-04921-8

Ⅰ.热… Ⅱ.俞… Ⅲ."大爆炸"宇宙学 Ⅳ.P159.3

中国版本图书馆 CIP 数据核字(2001)第 17255 号

书　　　名:热大爆炸宇宙学
著作责任者:俞允强　著
责 任 编 辑:周月梅　顾卫宇
标 准 书 号:ISBN 7-301-04921-8/O·464
出　版　者:北京大学出版社
地　　　址:北京市海淀区中关村北京大学校内　100871
网　　　址:http://cbs.pku.edu.cn/cbs.htm
电　　　话:出版部 62752015　发行部 62754140　编辑部 62752032
电 子 信 箱:zpup@pup.pku.edu.cn
印　刷　者:北京大学印刷厂印刷
发　行　者:北京大学出版社
经　销　者:新华书店
850×1168　32 开本　6.625 印张　172 千字
2001 年 4 月第 1 版　2001 年 4 月第 1 次印刷
定　　　价:14.00 元

《北京大学物理学丛书》
编委会名单

内容简介

本书清晰、翔实地介绍了宇宙学领域迄今为止取得的理论成果，其中包括对宇宙学所依赖的理论基础的阐明.宇宙大爆炸模型作为一个取得很多成功的理论设想，本书对此作了集中的探讨，包括模型所描绘的宇宙早期、宇宙演变的历史、3K微波背景辐射所给予理论可靠性的支持等，对模型所涉及的宇宙早期图像，如元素起源、核合成问题、暗物质问题、正反物质不对称问题，均作了仔细分析.最后对大爆炸宇宙学模型的成功和存在的问题作了总结，指出了宇宙学领域可能发展的契机和宇宙学模型有待继续改进之处.

本书叙述精确，论证确凿，行文周密而涵括丰富，可以作为物理学、天文学、宇宙学科研工作者和高等院校物理系师生获得宇宙学领域最新全面知识的读物.

前　言

物理学是自然科学的基础,是探讨物质结构和运动基本规律的前沿学科.几十年来,在生产技术发展的要求和推动下,人们对物理现象和物理学规律的探索研究不断取得新的突破.物理学的各分支学科有着突飞猛进的发展,丰富了人们对物质世界物理运动基本规律的认识和掌握,促进了许多和物理学紧密相关的交叉学科和技术学科的进步.物理学的发展是许多新兴学科、交叉学科和新技术学科产生、成长和发展的基础和前导.

为适应现代化建设的需要,为推动国内物理学的研究、提高物理教学水平,我们决定推出《北京大学物理学丛书》,请在物理学前沿进行科学研究和教学工作的著名物理学家和教授对现代物理学各分支领域的前沿发展做系统、全面的介绍,为广大物理学工作者和物理系的学生进一步开展物理学各分支领域的探索研究和学习,开展与物理学紧密相关的交叉学科和技术学科的研究和学习提供研究参考书、教学参考书和教材.

本丛书分两个层次.第一个层次是物理系本科生的基础课教材,这一教材系列,将在几十年来几代教师,特别是在北京大学教师的教学实践和教学经验积累的基础上,力求深入浅出、删繁就简,以适于全国大多数院校的物理系使用.它既吸收以往经典的物理教材的精华,尽可

能系统地、完整地、准确地讲解有关的物理学基本知识、基本概念、基本规律、基本方法;同时又注入科技发展的新观点和方法,介绍物理学的现代发展,使学生不仅能掌握物理学的基础知识,还能了解本学科的前沿课题和研究动向,提高学生的科学素质.第二个层次是研究生教材、研究生教学参考书和专题学术著作.这一系列将集中于一些发展迅速、已有开拓性进展、国际上活跃的学科方向和专题,介绍该学科方向的基本内容,力求充分反映该学科方向国内外前沿最新进展和研究成果.学术专著首先着眼于物理学的各分支学科,然后再扩展到与物理学紧密相关的交叉学科.

愿这套丛书的出版既能使国内著名物理学家和教授有机会将他们的累累硕果奉献给广大读者,又能对物理的教学和科学研究起到促进和推动作用.

《北京大学物理学丛书》编辑委员会

1997年3月

Preface

Physics is the foundation of natural sciences, a leading discipline of studying structures of matter and basic laws of motion. For several decades, driving by the demands of developing technology, the breakthrough in the studies of physical phenomenon and the laws of physics never end. During this period, all branches of Physics grew very fast and our knowledge of the basic laws governing the motion of the physical world was highly enriched. The growing of physics accelerated the progress of many physics related areas and technologies. The development of physics provided grounds and guidance for the birth and the growth of those new branches of physics, related areas and new technologies.

In order to catch up the main stream of the modernization and to give an impetus to scientific research and to improve teaching of physics in China. We decided to publish "*The Series of Advanced Physics of Peking University*". We invited those distinguished physicists and professors who worked in the frontier of physics to give series introductions to all branches of modern pbysics and recent developments in these fields. This series, as a consequence, provides textbooks and references for physicists and physics students in their studies of all branches of physics, related areas and technologies.

This series is divided into two sub-series of different levels, the first suh-series includes the textbooks of undergraduate

physics written by experienced teachers in Peking University in past decades. These textbooks were written concisely with deep insights and easier expressions, which adopt essences of physics textbook classics, explain fundamental concepts, laws and methods of physics in a systematic and rigorous way. In addition, these textbooks properly introduced tbe new approaches and the latest developments of physics for educational purposes. This sub-series is suitable for teaching of undergraduate physics for most universities and institutes in China. The second sub-series includes graduate textbook, references and academic writings. This sub-series focuses on the latest developments and accomplishments in the active subjects of relevant research with international interests and introductions to those of fast developing research fields. The topics of academic writings mainly cover all branches of physics, but it will be generalized to closely related areas.

We wisb tbe publication of this series could provide an opportunity for leading physicists and physics professors in China to show their fruitful accomplishments to general audience and to give an impetus to teacbing and research in physics.

Editorial board of
"The Series of Advanced Physics of Peking University"
March 1997

序

90 年代初，我曾应国家教委的物理教材编写指导组之约，撰写过一本“大爆炸宇宙学”．记得在教学和科研之余写了一整年，是在 1992 年春天完的稿．宇宙学是一门很年轻的学科，不断有新进展出现．当时我把有重要性的最新进展都写进去了．

稿件交给出版社后，不久就看到了清样．这很使我为做成了一件有益的事而高兴．可是这件事此后就石沉大海了．询问后被告知：书店的预定数太少，不能开印．就这样一拖三年多．1995 年夏天，在经过不应需要的交涉后，出版社才勉强准备开印 1000 册（后来实际上印了 1600 册）．出版后因发行渠道不通畅，常有认识或不认识的朋友给我来电话，询问此书何处有售．再过不算太久，书就售缺了．于是想读的人仍找不到这书，想满足读者需要的我无可奈何．这时我很为白花了力气而沮丧，但是没有勇气再交涉了．我不想去求出版社做它很不愿做的事．

2000 年春天，赵凯华老师建议我修订一下重新出版．我意识到仍有不少人需要这样的书，而且它经过 8 年也确实需要修订了，所以下决心做这件事．在把旧书看一遍后，我发现它已很不让我自己满意了．一方面是这 8 年宇宙学进展很多，有的还很重大．另一方面最我当初过于重视公式的来源和推导性的解释，不太适合于读者阅读理解．于是我明白，需要的不是增添一些新内容，而是得完全重写过．这意味着又要花一年左右的时间，但这是我乐意做的事．

宇宙学是一门典型的基础性学科．研究它的目的是满足好奇心．人们总想对大自然有更多一点的了解和更深一点的理解．就像人人想知道物质的最小基元是什么一样，也想知道最大的宇宙是

什么样子的. 人人都应该知道,例如化学元素是原子核决定的,原子核内有质子和中子,质子及中子是夸克组成的等. 其实人人也应知道,宇宙可看成星系组成的气体,这气体中还有温度为 3 K 的辐射场,最早的化学元素是在宇宙年龄为 3 分钟时出现的等. 但是实际上,宇宙常识的普及远不如微观知识的普及. 后者早已进入中学教育,而对前者,连学物理或天文的大学生都知之甚少. 作为一个宇宙学专业工作者,我觉得有责任来改变这种状况.

就这本小书的性质而言,它是算不上一本通俗读物的. 我是为具有一些大学物理知识而想了解宇宙的读者写的. 我对自己提的要求是:(一) 尽量全面一些,把被研究过的宇宙学问题[①]都写进去;(二) 写得深入一些,使读者不像在听故事,而能用自己的物理知识理解它.(三) 写得客观一些,有的结果已很肯定,有的很初步甚至是猜测性的,不能误导[②]读者. 此外我心目中还有一个动机:希望通过此书吸引更多的年轻人对宇宙学感兴趣. 我国的宇宙学研究队伍太小了,平均水平也低. 事情远不该是这样的. 因此我也想用这本书为愿意进入宇宙学研究的读者做入门的导引. 我写下这些想法,是要请读者给我反馈:效果与愿望是否能一致.

最后我愿向北京大学出版社表示敬意. 副总编周月梅和书记王明舟都对我说过,他们的方针是用畅销书赢得的经济效益来"养"不可能大量畅销的科技书. 这确实不是短视的出版者肯做和敢做的事. 他们是在为我国的科学事业做着一件好事.

俞允强

2000 年 12 月于北京大学

① 实际上除宇宙创生问题外,其他大问题都讨论到了.

② 若对宇宙学的观测结果或理论想法采取"见风即雨"的态度,极易引伸出误导性的结论. 在我看来,轻易叫人相信,或轻易提出"挑战"都属于误导.

目　录

引　言

宇宙指的是物质世界的一切.它是人类所面对的最巨大的客体.正是这一特点使它历来吸引着各种人的兴趣.诗人对它抒发情怀,哲人借它阐发哲理.可是近代的宇宙学家却是把它当作一个普通的物理对象来研究的.我们图的是对它有一个具体而真实的了解.

初一想,地球上的物质已十分纷杂,这包罗一切的巨大对象如何研究?我们自然知道,地球上的一切在宇宙中只是细微末节.为研究宇宙,须放眼于地球之外.抬头看天,见到的是满天星斗,人们都叫它们为星星.实际上只有那些离我们较近的光点才是恒星.远的光点不是恒星,而是很大的恒星集团,叫星系.这些肉眼能见的和望远镜中才能见的,还有连用望远镜也不能见的天体,才是宇宙物质.

宇宙学的目标不是研究这些天体本身,这是它与天文学或天体物理学的区别.宇宙学是借助各种天体带来的信息,以研究宇宙的整体状况和行为.这目标决定了它的视野十分宽广.宇宙中可能被观测部分的大小超过百亿光年.整个宇宙或许无限,也或许有限.即使它为有限,其大小比百亿光年肯定要大很多很多.空间上的广袤是宇宙的特点之一,也是研究它的困难根源之一.我们不仅必须能观测非常远处的天体,而且还须有能力从观测中析取可靠而定量的信息.人类起步研究宇宙并不晚,大约是在20世纪初.它的蓬勃发展却仅是近30年的事.后来靠的首先是天文观测上的进步.

宇宙不仅在空间横向地展开它的复杂性,还在时间纵深上展开了复杂性.由于光的传播需要时间,所以对一亿光年远的天体,

看到的是它在约一亿年前的面貌，而不是今天的面貌. 这是一个很普通的道理，但也是很深刻的道理. 这样看来，宇宙在空间上的复杂性和时间上的复杂性只是同一件事的两个侧面. 天文学家越往远处观测，他们所看到的是越年轻的宇宙.

人们从 20 世纪 20 年代起就知道了宇宙不是不变的，而是在膨胀. 自然界所有看来像亘古不变的东西实际上都有缓慢的变化. 这种变化叫演化. 生物物种在演化，恒星和星系在演化，宇宙也不例外地在演化. 从前远古的宇宙是什么样子的?它如何一步步地演化成了今天的模样? 这是宇宙学想弄清楚的主要内容.

现在知道，我们的宇宙已演化了 150 亿年左右. 有趣的是，过去的宇宙完全不是今天宇宙的缩影. 它们之间有着质的差别. 年轻的宇宙中曾没有恒星，也没有星系. 它曾只是一片炽热的气体. 更令人惊讶的是，在更年轻的宇宙中没有碳、氢、氧、氮，等等，连化学元素这样基本的物质形态也是在宇宙演化中产生的. 这些结果其实都非常可信，但又非常难以令人置信. 宇宙学的多彩和魅力正在这种地方.

当我们按已知的物理规律推断出一段段遥远的宇宙历史，怎么证明它是真实而可信的?宇宙学必须对此作出回答. 实际上这才是要研究的核心. 今天的宇宙学是作为物理学的一个分支而发展的. 因此回答的办法只有一个，那就是提供直接或间接的事实证据. 宇宙演化史感兴趣的至少是十亿年前的事. 要找出证据可真谈何容易! 所以，今天的宇宙学能理直气壮、振振有词地谈论宇宙最初几万年的事(3K 微波背景辐射)，甚至最初 1 小时的事(原初核合成)，实在是魅力无穷，令人在心灵上感受到震撼.

在 3K 背景辐射发现以来的 30 余年中，热大爆炸宇宙理论已赢得了无可争议的成功. 这是事实，但是成功的事实不要掩盖其弱点. 瑕不掩瑜，瑜也不掩瑕. 总体说来，宇宙学的发展还处于它的未成年期. 尚不能回答的问题远比已有答案的问题多. 已能回答的问题中，答案的牢靠程度也各各不同. 这才是今天宇宙学的真实面

貌.因此要把这副面貌客观而全面地描绘出来也不是易事.

本书显然想把宇宙学中闪光的一面展示出来.它不仅指那些已找到了证据的重要结果,而且还包括着许多尚未找到证据的非常杰出的新思想.例如真空有能量的思想,非重子是宇宙主体的思想等.这些新思想很可能反映宇宙学的将来,也可能是诱使物理学发展的生长点.其重要性不亚于那些已有答案的问题.但是对此必须非常小心.在找到可靠证据前,再杰出的思想也仅是待证实的猜想.不能把它与事实混同.总而言之,把宇宙学现状中的成功和不足两面都展示出来,才是本书的全部目的.

第一章　宇宙的面貌

宇宙是自然界广袤最大、历史最久的客体.要研究它的发展过程,只能先从认识它的面貌入手.越来越深入的天文观测,使人们看到宇宙的结构很复杂.但在深入理解了其复杂性后,人们又察觉了它背后的简单性:在宇观尺度上,今天整个宇宙可看成一大片膨胀着的均匀介质,而我们的银河系只是这介质中的一颗"分子".

一、宇宙级距离的测量

我们生活在地球上.这是一个小小的,在空中飞行的固体球.我们对它本身比较熟悉.地球之外的整个世界是什么样子?这是上千年来人类智慧总想冲击的一个难题.

经哥白尼、伽里略和牛顿等一批先驱者的努力,人们至17世纪才弄清:地球周围有一个太阳系,供给地球以光明的太阳是它的中心,而地球只是其中一颗普通的行星.太阳系广延达10^{13} m.而今天的观测宇宙的广延是10^{26} m.在尺度上相差13个量级!太阳系与观测宇宙相比,大致相当与把一粒尘埃与太阳相比.太阳系在宇宙中是非常微不足道的.因此,为了认识宇宙的面貌,还需要大大地扩展我们的视野.

太阳系之外,天穹上还有满天的星斗.这是自然界中最壮观,也最为人熟悉的景象之一.可是,我们不能直觉地感受不同星斗的

远近，因而这表观景象仅是二维的. 它不能体现宇宙的立体面貌. 为知道满天星斗的立体分布，必须先学会测量遥远天体和我们的距离. 这可不是一个简单的任务. 实际上，人们很长时间来没能了解宇宙的结构，正是这个难题挡了道.

地面上常用的大距离测距法是三角形法(图 1.1(a)). 18 世纪中叶，开始有人用这方法测量了月地间的距离. 很明显，用三角形法测量的距离越大，需要引用的基线也越长. 为测量恒星间的距离，用整个地球直径做基线都嫌太短. 于是人们想到用地球绕太阳转动的轨道直径做基线. 这确是一个很聪明的办法. 在地球轨道平面上总能找到一条直径，使太阳与所测恒星的连线正好与它垂直(图 1.1(b)). 这样，测出这恒星对地球轨道**半径**的张角 α，就能用已知的日地距(即轨道半径)来推出该恒星的距离了. 当 $\alpha=1''$，算出相应的距离是 3.1×10^{16} m. 天文上常把这个量当作距离单位，称 1 秒差距，记作 1 pc. 它与光年的关系为 1 pc＝3.3 光年. 如测到某恒星的张角 $\alpha=0.1''$，则它距太阳 10 pc，即 33 光年.

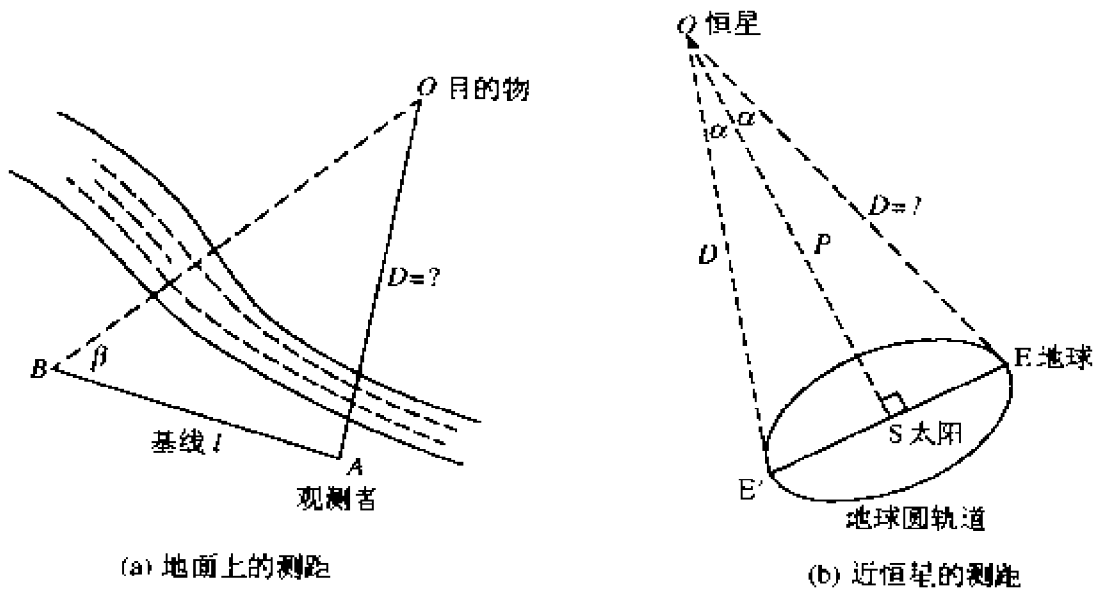


图 1.1　三角形测距法

有了这方法，人们才知道，离太阳最近的恒星是半人马座的 α 星，其距离为 1.3 pc. 全天穹表观最明亮的天狼星距我们 2.67 pc，是第七个最近的恒星. 对太远的恒星，会因其张角 α 太小而测不出距离. 三角测距法在天文上的有效范围约是 10^2 pc. 这使人类的视野从太阳系的尺度拓宽了 5 个量级. 但是在浩瀚的宇宙中，我们看清了的仍只是沧海一粟. 具体讲，在这尺度范围内看到的是稀疏地分布着几千颗恒星. 它只是我们银河系中很小的一部分，而还远不是宇宙整体面貌的一个典型代表. 因此，要认识宇宙，我们还须去寻找更有力的测距方法.

若把天体在单位时间内辐射的光能用光度 L 描述，而把肉眼或仪器感受到的亮暗程度，即单位时间在单位面积上接受到的光能叫视亮度 B，那么视亮度不仅取决于光源的光度，还取决于距离（图 1.2），即

$$B = L/(4\pi d^2), \tag{1.1}$$

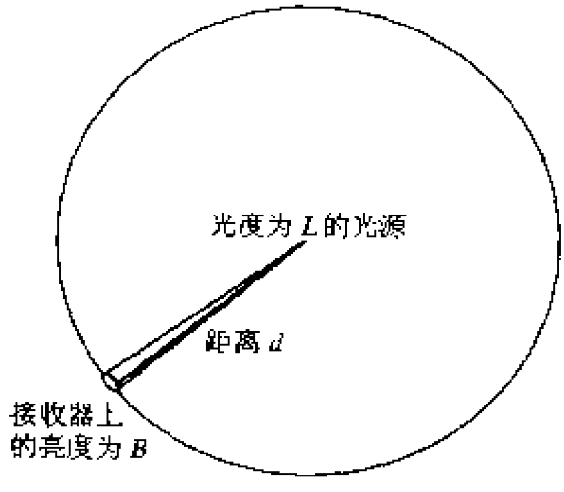


图 1.2　视亮度与光度的关系

其中 d 是光源与观测者的距离. 天体的视亮度 B 较容易直接测量. 这样，若已测定某恒星的距离，就能推知其光度. 以几千个已测定距离的恒星为基础，人们就知道了一批恒星的光度. 由此发现，不同恒星的光度差别很大，而太阳是一颗中等偏暗的恒星. 关系式

(1.1)也反过来给我们以启示：若可能事先知道某天体的光度，那么通过测量它的视亮度，就能推知其距离.这正是用天体物理(指非几何)方法测距的基本思想.问题在于：如何在测定其距离之前能先知道它的光度？在20世纪头十几年里，人们就发明了两种方法.

图1.3中以恒星的表面温度(或光谱型)为横轴，光度为纵轴.对于这两个量都已测定的恒星就在坐标面上有一个代表点.在对大量恒星作统计研究时发现，90%以上的恒星的代表点落在一条带上(见图1.3).这条带被称为恒星的主序列，相应的恒星叫主序星.这经验规律告诉我们，主序星的光度与其表面温度有确定的关联.恒星的表面温度较容易测定.按这经验规律，在测到它的表面温度后，就可以推知其光度.那么再进一步结合它的视亮度，这颗恒星的距离就确定了.用这方法，银河系中大多数恒星的距离都能确定下来.

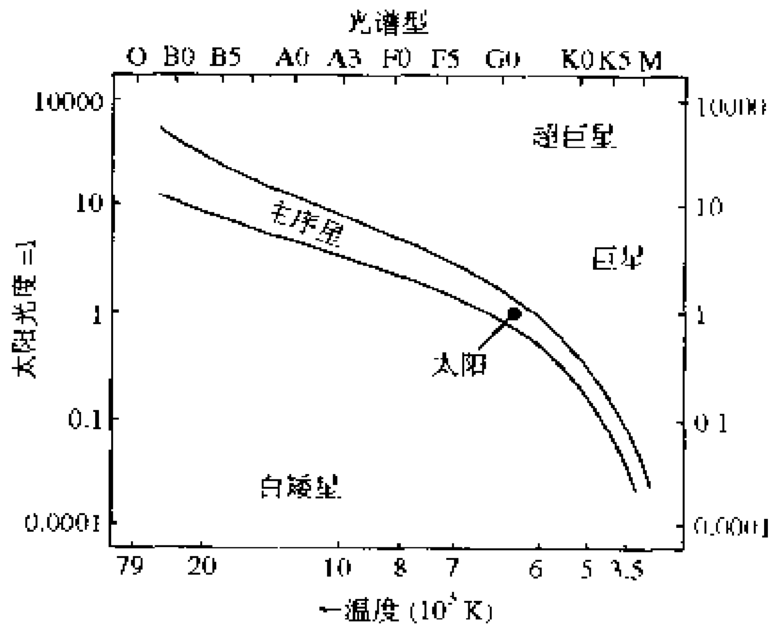
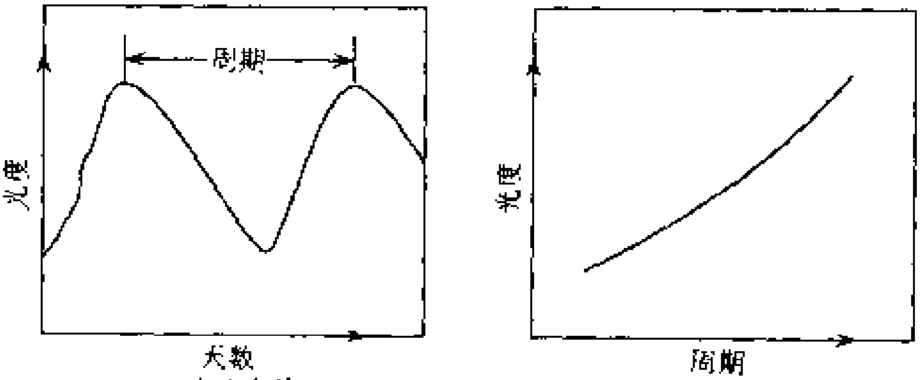


图1.3 赫兹勃龙-罗素图

主序星是以内部氢聚变为能源的恒星，它是恒星演化的第一

阶段.在演化的晚期,大量恒星的星体会发生周期性的膨胀和收缩,其光度和亮度也相应有周期变化(图 1.4(a)).这种星叫周期变星.1912 年,Leavitt 在研究麦哲伦云(银河系最邻近的星系)中的造父变星(周期变星中的一类)时发现,不同变星的光变周期与其光度有确定的关联,如图 1.4(b)所示意.变星的光变周期较容易测定.这样,用光变周期推断其光度,再由光度来推断距离,就又构成了一种测距方法.



(a) 光变曲线

(b) 周期-光度关系

图 1.4　用造父变星测距

用周期变星测距有两个好处.若在远处的星系中找到了周期变星,就能用它的光变周期来测定其距离.由于星系的大小比这距离小很多,因此所定出的实际上就是该星系和我们的距离.在这意义上,人们把这种变星叫星系的距离指示器.这是一个好处.另一好处是造父变星的光度通常很大,约比太阳亮 5 个量级,因此能用它来测量较远的星系的距离.在今天的技术下,能测量的距离达到了 10 Mpc 的范围.与几何测距法相比,它把我们的测距能力又提高了 5 个量级.

当然,10 Mpc 的范围与宇宙相比还是太小.为此,后来的人们又进一步发明了许多其他方法.我们不一一讨论了.总的精神仍然一样.在已能测定距离的范围内寻找天体光度和其他较容易观测的量之间的联系的经验规律.然后把这经验规律用于更远的星系.

先间接地确定它的光度，再进一步推断它和我们的距离. 随着测距能力的逐步提高，今天人们对整个观测宇宙的面貌已有较清楚的了解了. 这里也值得指出，至今用天体物理方法测距的精确程度尚难以很高. 现在对远处星系的距离的测定大约有一倍的误差. 后面将看到，这是造成宇宙学信息不确定性的一个重要原因.

二、层次性的结构

随着测距能力的逐步提高，人们逐渐在越来越大的尺度上对宇宙物质的分布有了立体的观念. 这里第一个重要进展是认识了银河系. 它包含着两重含义：一是了解了银河系的形状，即银河系内的物质分布；二是认识了银河之外还有星系存在.

银河系是太阳所属的一个庞大的恒星集团. 它大约包含有 10^{11}颗恒星. 现在把这样的恒星集团统称为星系. 银河系的概貌如图 1.5 所示. 它的大部分恒星集中在一个扁平的区域内，被称为银盘. 盘的直径为 25 kpc，厚度约为 2 kpc. 盘的中心有一球状隆起，被称为核球. 若鸟瞰这银盘，它是由从核球向外延伸的几条旋臂所组成. 太阳位于一条旋臂上，距银河系中心约有 8 kpc. 银盘上下有球状的晕区，叫银晕，其中的恒星较稀疏. 晕区的直径为 30 kpc，晕内物质的质量约只占银河系总质量的 10%. 除了恒星外，星际稀薄气体也是银河系的重要组分. 这就是我们所在的星系的概貌. 太阳作为我们的恒星，其质量、光度或位置都没有太大的特殊性，而只是银河系内的普通一员. 这正是哥白尼思想在星系尺度上的体现.

在认识了银河系之后，重要的问题是：是否天穹上一切光点都是银河系的一部分？这问题在历史上有过很多哲理的争论，但是实际上它只有用观测才能回答. 要通过观测来给出回答并不容易. 设想有一个类似于银河系的恒星集团处于 500 kpc 的距离上（记得银河系自身大小仅为 30 kpc），其表观亮度与 2 kpc 处一个类似

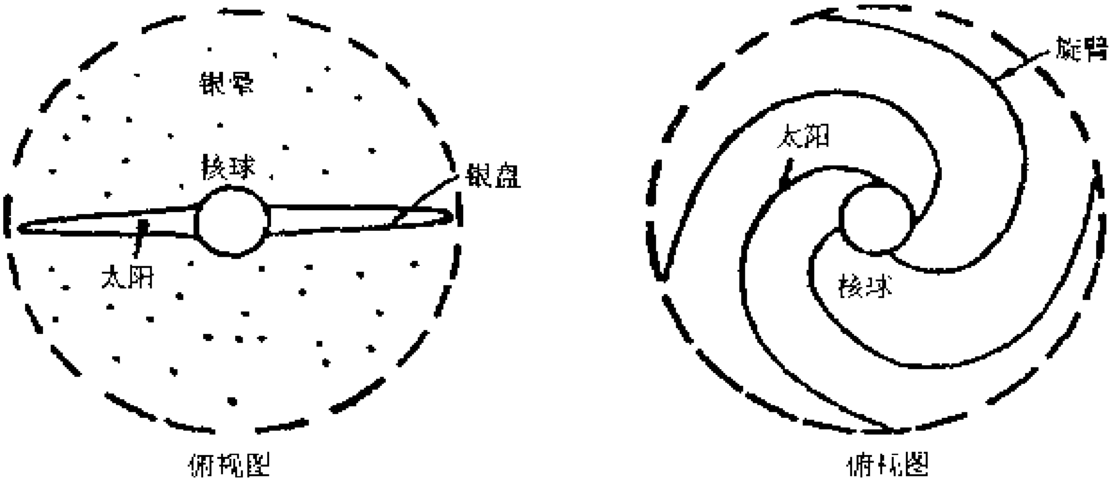


图 1.5 银河系的结构

太阳的恒星是一样的.因此,为判断天穹上的某个光点是银河系内的恒星,或银河外的恒星集团,必须要测量它的距离.在当时,除了属于银河系的恒星外,已经知道还有不少云块状的天体,它们被笼统地称为星云.经过很多年的争论后,有人在某些星云中找到了造父变星.用变星方法测定其距离,才知道这些被叫做星云的天体其实不是恒星际的云块,而是远在银河系之外的很庞大的恒星集团.这发现第一次使人们知道:银河系仅是宇宙的一部分,而不是全部.这无疑是人们认识宇宙的一大进步.这样的恒星集团很快被发现了很多,人们才建立了星系的概念.星系的普遍存在,表明它是宇宙结构中比恒星更大的一个新的层次.

不同星系的质量差别很大.银河系的质量约为 $10^{11}\ M_{\odot}$(太阳质量单位).在较容易观测到的明亮星系中,这是中等的、典型的大小.随着望远镜的观测能力越来越大,天文学家发现了许多很暗的小星系.它们的质量可低达 $10^{6}\ M_{\odot}$.因此银河系实际上属于大型的星系.不同星系的外形也很不相同.有旋涡状(银河系是一例)、椭球状和不规则状的.通过很大范围的观测巡视,人们认识到,宇宙空间中弥散地分布着无数大小不同和形状不同的星系.因此概括地讲,宇宙介质可以看成是以星系为"分子"所组成的"气体".

在知道宇宙介质可看成由星系组成的气体后，接着的重要问题是：星系在宇宙空间的分布是随机的、无序的吗？如果答案是肯定的，那么这气体自然是均匀的，但天文观测的回答是不完全如此，星系的分布仍有结团性. 其实考虑到万有引力的存在，这在物理上是很可理解的事. 观测表明，星系的结团性远比恒星弱. 大约只有 10%的星系从属于由几百、上千个星系结合成的大型集团，叫星系团. 大部分星系只结合成含十几、几十或上百个成员的小集团，它们被叫作星系群. 还有不少星系只是三三两两地结团. 无论如何，由星系结合而成的集团代表了宇宙结构中又一个更大的层次. 我们可以笼统地称它为星系团的层次，它的尺度为 Mpc 的量级.

那么还有更大层次的物质结团吗？在尺度为 100 Mpc 的范围内的星系个数达 10^6. 须逐个测定其距离，才能知道它们的立体分布，因此要用观测来澄清这问题是很艰巨的任务. 总究在 70 年代后期，人们发现了超星系团（简称超团）的存在. 在超团中通常只包含几个大型的星系团，而大部分是很小或极小的星系集团，因此超团的总质量仅比大的星系团多一个量级. 超团的外形不呈球状，而是扁长形的. 到 80 年代末，人们又发现了相应尺度的空洞，即在很大（例如几十 Mpc）范围内几乎没有星系存在的区域. 现在已观测到最大空洞的直径达 50 Mpc. 按目前的认识，超星系团和空洞代表了宇宙物质分布上的最大层次. 但是对宇宙在 50 Mpc 以上是否还有新的结构，这问题依然有人在研究.

到这里让我们最后问：整个宇宙总共有多大？这问题看来简单，却不是能简单地回答的. 事情的一面是由于宇宙空间的弯曲性，它的总体积可能有限，也可能无限. 但事情还有另一面. 由于今天的宇宙年龄 t_0 是有限大，光在这时间内能传播的距离（ct_0 的量级）也有限. 因此，不管浩瀚宇宙的总大小为有限或无限，我们能观测的部分却总是有限的. 这些问题都将在后面讨论到. 这里仅想指

出：可观测宇宙的大小[1]约是 10^4 Mpc，它比超星系团的尺度仅大三个量级左右．这样看来，上面谈到的认识已大体反映宇宙的全貌了．

总之，从宇宙学的眼光看，我们宜把星系当作宇宙物质的基本单元，而把恒星只看作星系内部的细节．至于太阳系则是更细的细节，完全不属于宇宙学关心的范围了．宇观地看宇宙，它具有层次性的结构．把星系当基本层次，其上还有星系团的层次和超团及空洞的层次．这就是今天对宇宙面貌的基本认识．

三、宇宙学原理

当我们对任何客体作物理的理论研究时，总必须先把它简化．实际客体都是综合的、复杂的．它包含许多并非本质的次要属性．若把一切属性都考虑在内，就难以对它作深入的研究了．简化的要领是把客体的次要属性丢掉，但与此同时，必须把客体的主要属性保留下来．这样，人们所研究的虽是客体的简化模型，而模型的主要行为却与真实客体是一致，或相接近的．若没有把主要属性抓住，模型的基本行为将也不反映客体的实际，那么研究这模型就没有意义了．这是做基本理论研究的一般方法．人们在研究宇宙时也是这样做的．

第一个物理的宇宙模型是爱因斯坦在 1917 年提出的．当时人们对宇宙的整体面貌还完全没有了解．因此他主要凭直觉，作了两个简化假设：(一) 宇宙物质在空间上是均匀和各向同性的．今天人们把它叫做宇宙学原理．(二) 宇宙物质的分布是不随时间变化的．在这两条假设基础上，他建立了静态的宇宙模型．20 年代初，天文学家初步发现了宇宙在膨胀的迹象．敏感的人很快意识到，爱

---

[1] 这指按广义相对论定义的固有距离，它与测量上得到的(光度)距离是有区别的．

因斯坦的后一条假设错了. Friedmann 在只保留前一假设的基础上建立了膨胀的宇宙模型. 它就是今天的标准宇宙模型的雏形. 由此，我们必须对宇宙学原理与真实宇宙的关系作出讨论.

如上所述，宇宙学原理原本只是爱因斯坦的猜想，而并没有事实根据. 后人沿用了下来，很大程度上是把它当作一个方便的工作假设. 因为如果宇宙不是均匀的，那么研究宇宙动力学的条件就更不成熟了. 当然，均匀假设到底是否符合事实是根本问题. 膨胀的宇宙理论在 60 年代前一直得不到学术界的认可，这基本假设没有证据是重要原因. 这方面证据的获得是 60 年代后的事.

一片气体被认为是均匀的，意指不同地方的同大体元内有同样的质量. 但是气体有微观结构，即它是由分子所组成. 因此，若把体元大小取得与分子可比拟，不同处体元内的质量是不会相等的. 这道理说明"均匀"是一个宏观概念. 现在需要把这概念用到宇宙学中来. 我们已经知道，在星系团或超星系团的尺度以下，宇宙物质的分布是结团性的. 因此，它的均匀性应是一个**宇观**概念.

上节中已指出，宇宙可看成由星系为"分子"的气体. 由于星系间有长程力，即引力作用，它使得星系有结团行为. 因此，若以星系的平均间距(约 1 Mpc)为体元尺度，不同地方体元内的质量必会有显著的不同. 当把体元尺度放大到 10 Mpc，即星系团与超星系团之间的大小，可以预料，体元内的质量仍会有不小的差别. 宇宙介质的均匀性应表现为在更大的尺度上偏离均匀的程度很小，且随尺度的加大而越来越小. 为了能用实测来证实或否定这概念，我们可把不均匀程度定量化.

考虑直径为 λ 的球形区. 把宇宙的不均匀性完全抹平后，这球区内的质量为 M. 在不严格均匀的实际宇宙中，这区内的质量记 $M+\delta M$，δM 反映它对均匀背景的偏离. 显然不同地方有不同的偏离大小. 对全宇宙讲，δM 的平均值必定为零. 有的地方球内的质量超过平均值，必有别的地方低于平均值. 因此我们应把各处的 δM 平方后再作平均，得到 $\langle\delta M^2\rangle$，它表示整个宇宙偏离均匀的程

度. 按上面的分析,这样定义的偏离程度与尺度 λ 的大小有关. 原则上,这 $\langle \delta M^2 \rangle$ 值随 λ 的变化可以通过对大范围巡天资料的分析来得到. 如果实际情况显示出 $\langle \delta M^2 \rangle$ 值随 λ 的增大面减小,这就是宇观地均匀的证据.

当然作这种分析是很艰难的任务,但是初步结果已有了. 首先人们发现,使平均相对偏离 $\langle (\delta M/M)^2 \rangle = 1$ 的尺度是 8 Mpc. 在更大尺度上,相对偏离是小于 1 的. 1990 年有一个研究组用 IRAS 科学卫星得到的巡天资料作分析,得出的结果是: $\langle (\delta M/M)^2 \rangle$ 在 15 Mpc 的尺度上为 0.18,在 45 Mpc 的尺度上为 0.067. 这结果确实告诉我们: 尺度取得越大,宇宙气体偏离均匀程度越小. 另外还表明,在 50 Mpc 以上,平均偏离已小到 25%以下. 虽然这类分析在细节上还颇可争议,但无论如何,它们对宇宙学原理给出了有力的支持.

更强烈的证据来自早期宇宙. 按宇宙学理论,今天存在的星系是在宇宙演化中产生的. 在一切星系形成前,宇宙介质是高度均匀的普通气体(参看第四章). 3K 微波背景辐射产生于一切星系尚远没有形成前. 那时不同地方的密度不均匀,将通过热辐射温度的各向异性表现出来,因面它是可以测量的. 这问题将在第五章中讨论. 这里仅指出: 微波背景辐射温度的观测,证实了当时宇宙密度的不均匀程度仅为 10^{-3}. 这是对宇宙学原理的更有力的支持. 总之,宇宙学原理作为从理论上研究宇宙的出发点,现在已不再是一个猜想性的或图方便的假设. 事实已证实,它是实际宇宙的一个合理的近似描写.

四、宇宙的膨胀

要对宇宙作理论研究,总需要以一些基本的观测事实做基础. 上面谈到的宇宙均匀性是一方面重要事实. 另一方面重要的事实是宇宙在膨胀.

在爱因斯坦解出他的静态宇宙模型时，天文学家 Slipher 正在观测远处旋涡状星云的光谱波长的移动. 他发现，多数星云发出的光谱在我们看来有红移，仅少数有蓝移. 按当时的物理知识，这种波长的变化只能归因于光源运动的多普勒效应. 那么，多数星云的光谱有红移，表明它们都在向远离我们的方向退行. 这是最早被人们发现的宇宙膨胀的迹象.

Slipher 的发现发表在 1923 年. 当时还不知道这些星云是银河系之外的天体. 同年，另一位天文学家哈勃可靠地用观测论证了某些旋涡星云是远在银河系之外，而与银河系相类似的恒星集团. 若在银河之外的大量星系都在向远离我们的方向运动，那么推而广之，宇宙作为星系组成的气体，它在膨胀的图象就形成了. 当时有好几位理论家如 Eddington 和 Friedmann 等都意识到了这一点. 有人把这事实和看法告诉了爱因斯坦，他也立即意识到他的静态宇宙假设应当抛弃.

由于哈勃已掌握了一些测量星系距离的方法，他进一步研究了星系的退行红移 z 与距离 R 的关系. 当时观测到的红移都很小，按非相对论的多普勒效应公式，退行红移与退行速度成正比，即 $z=v/c$，其中 c 是光速. 他在 1929 年经验地发现，不同星系的退行速度与它的距离成正比. 后人把这规律叫哈勃定律，写作

$$v = H_0 R, \tag{1.2}$$

其中的比例系数 H_0 被称为哈勃常数.

让我们注意，哈勃常数作为速率与距离之比，它所反映的是宇宙的线膨胀率. 容易论证，线膨胀率为常数，表明宇宙中不同地方的密度变率是一样的. 图 1.6 中直观地画出了这样的论证. 在均匀的宇宙介质中取出四个等距分布的星系 O（我们的位置）、A_1、A_2 和 A_3. 按哈勃定律，从 O 看另三个星系，它们的膨胀速度比为 $v_1 : v_2 : v_3 = 1 : 2 : 3$. 当 A_1 达到 A_1' 时，A_2 和 A_3 分别达到 A_2' 和 A_3'. 这样易于看出，O、A_1'、A_2' 和 A_3' 仍是四个等距分布的星系. 于是这就说明，原来均匀的气体，在按哈勃定律膨胀后依然是均匀

的. 因此哈勃定律不仅告诉我们宇宙在膨胀，而且暗示了宇宙在它的演化历史上一直是均匀的.

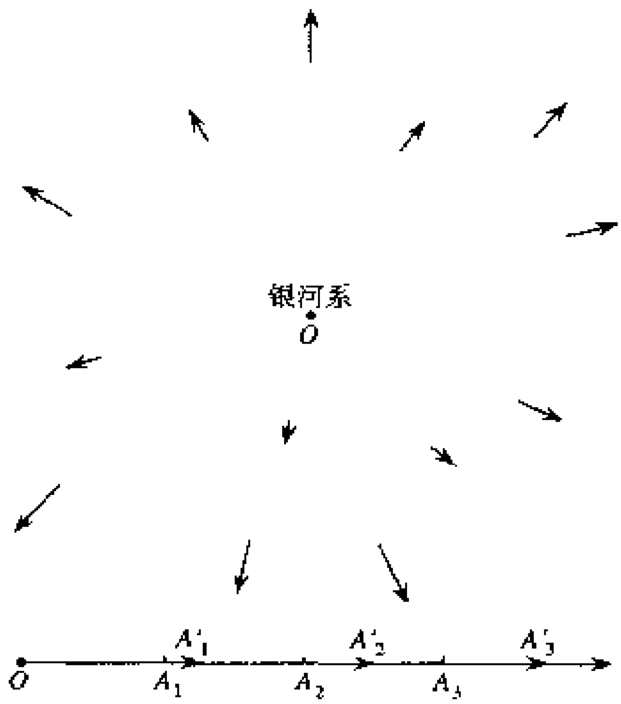


图 1.6 按哈勃定律的膨胀

哈勃定律表明一切星系都以我们为中心向四方散开. 这容易给人以错觉，好像我们处于宇宙的中心. 在上面论证的基础上，这错觉就容易澄清了. 对永远均匀的宇宙讲，空间不同位置是完全平等的，我们所处的位置不可能有任何特殊性. 如果其他星系中有观测者，那么他会发现，宇宙相对他的星系也是按哈勃定律膨胀的. 人们常把这一观念叫宇宙学的哥白尼原理.

确切说来，哈勃常数所反映的是**今天**宇宙介质的膨胀速率. 这是一个非常重要的宇宙学参量. 让我们注意一下它的大小.

用 Mpc 作为距离 R 的单位，用 km/s 作为速度的单位，则 H_0 的单位是 $\mathrm{km\cdot s^{-1}\cdot Mpc^{-1}}$. 哈勃本人在 1931 年首次定出的值是

$$H_0 = 500\ \mathrm{km\cdot s^{-1}\cdot Mpc^{-1}}, \tag{1.3}$$

即距离为 1 Mpc 的星系的膨胀速率为 500 km/s. 正是这过大的

膨胀率，导致了宇宙年龄过小（2 Ga①以下）的预言（见第二章）. 它曾使膨胀宇宙理论失去了当时人们的信任. 到 50 年代，后一代天文学家才发现哈勃在距离测定上有错误（结果显著偏小），并弄清了 H_0 的实际值要小得多. 但是至今用不同方法测量远处星系的距离依然有不小的差别，因此相应定出的哈勃常数值仍相当弥散. 它是

$$H_0 = 40\text{--}100\ \mathrm{km \cdot s^{-1} \cdot Mpc^{-1}}. \tag{1.4}$$

经过几十年的努力，包括发射哈勃空间望远镜到大气外去测量，结果 H_0 到底多大的问题还没有解决. 按专家们现在的共识，它的实际值应在 50 到 80 之间. 这遗留下来的不确定性给宇宙学研究带来很多不利，后面我们将屡屡看到.

关于哈勃常数，还有一点值得注意的细节. 它的倒数具有时间的量纲. 若接受 $H_0 = 50\text{--}80\ \mathrm{km \cdot s^{-1} \cdot Mpc^{-1}}$，利用 $1\ \mathrm{Mpc} = 3.1 \times 10^{19}\ \mathrm{km}$，易于折合出

$$H_0^{-1} = (3.9\text{--}6.2) \times 10^{17}\ \mathrm{s} = (1.3\text{--}2.0) \times 10^{10}\ \mathrm{a}. \tag{1.5}$$

以后将从理论上论证，它与宇宙年龄 t_0 是同量级的. 因此在做简单讨论时，人们常把 H_0^{-1} 当作宇宙年龄. 其实更确切些说，它是宇宙年龄的上限. 此外按同样道理，cH_0^{-1} 应是观测宇宙大小的量级. 相应的值是

$$cH_0^{-1} = 3800\text{--}6000\ \mathrm{Mpc}. \tag{1.6}$$

本节开头时已指出，哈勃定律作为经验规律是在红移 z 很小时得到的. 从(1.2)式看，并注意到 $z = v/c$，红移很小意味着该星系与我们的距离须远小于 cH_0^{-1}（即观测宇宙的大小）. 对于远的星系，测量表明其光频的红移会超过 1，但这情况下形式如(1.2)的哈勃定律是不适用的. 这时红移与距离的关系是什么，那既是实测须回答的问题，也是宇宙学理论应回答的问题. 问题的另一面

① a 是 anno 的缩写，代表年. $1\ \mathrm{Ga} = 10^9\ \mathrm{a}$.

是，哈勃定律对太近的星系也不适用.我们观测到的星系运动包括两部分.一部分是宇宙膨胀带来的，另一部分是它自身的无规运动.后者类似于气体分子的热运动，天文上称它为星系的本动.它的大小约是每秒几百公里.只有当膨胀引起的速度显著地超过本动，以致后者可忽略，才会有(1.2)形式的膨胀规律出现.这就要求被测星系与我们的距离超过 10 Mpc. 否则为知道膨胀速度，必须在测到的速度中扣除本动，这又是一个困难的任务.因此，尽管较近星系的距离较容易测准，但我们并不容易用它来准确地定出哈勃常数.

第二章 标准宇宙模型

在观测证实可把宇宙看成一片始终均匀的介质后，再假定宇宙的行为服从广义相对论，那么它的膨胀动力学方程就具体地确定了．我们的目标是要把宇宙膨胀的整个历史确定下来．为此，还必须先通过观测来获得一组“初”条件．这就是宇宙理论的一个标准的框架．当然，它的正确与否是需要用进一步的测量来检验的．

一、牛顿引力定律的失效

从力学上讲，宇宙是一个自引力作用下的系统．人们自然希望以牛顿第二定律和牛顿引力定律为基础来研究它的膨胀过程．可惜在这一问题面前，牛顿力学不能构成一个合适的研究框架．

在牛顿理论中，空间是独立于运动着的物质之外的背景．它是平坦的，从而也必定是无限的．按宇宙学原理，宇宙物质均匀地充满了全空间，所以物质也必定是无限的．当人们试图写出这无限介质中一个任意质元所受到的引力时，发现牛顿引力理论完全失效．

图 2.1 画出了来自某小锥体中物质的引力（对质元 O）．把锥体分成无数个厚度为 dr 的薄壳．考虑距 O 为 r 的壳．牛顿万有引力定律讲，引力与距离平方反比，而与壳内的质量正比．注意壳内的质量又正比于 r^2，因此远近不同的壳对质元 O 产生同样大的引力．这么，这小锥体中物质对 O 的总引力是无穷大！ 当然，对顶方向上的物质将在相反方向上也产生一个无穷大的引力．那么两个

无穷大之差是多大？这在数学上是没有明确答案的.

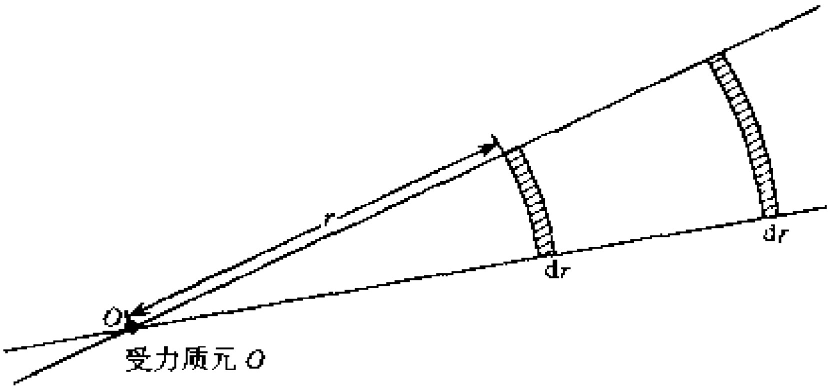


图 2.1　来自一个小锥内的引力

如果宇宙介质是理想地均匀，那么两者的合力将是零. 但是我们的宇宙，如同一切物体一样，并不会是理想地均匀的. 于是为了得到合力，我们必须先知道不均匀的全部细节. 而最终得到的结果则可能是零，可能是有限数，也可能还是无穷大. 人们把这个古怪的结果叫 Seelinger 佯谬.

如果能算出所受的引力，则可以用牛顿第二定律来求它的运动. 现在发现引力写不出来，当然这样就做不下去了. 那么牛顿的万有引力定律为什么会失效？

为了好理解，让我们把引力与电力作对比. 库仑定律与牛顿引力定律有相同的形式，但前者只是静止电荷产生电力的规律. 当电荷在运动，它不仅产生电力，还会产生磁力，而且其中电力也不再服从库仑定律了. 这也就是说，库仑定律不是电力的普遍规律. 同样，牛顿的引力定律也不是引力的普遍规律. 它也只是物体静止时产生引力的规律. 宇宙介质并不静止，而是在膨胀. 或许 Seelinger 佯谬的出现是偶然的，但牛顿定律对宇宙动力学不适用却是必然的.

我们知道，电磁场的一般规律是麦克斯韦的理论. 为了研究宇宙膨胀的动力学，我们需要用引力场的一般理论. 这就是爱因斯坦

的广义相对论.宇宙学理论框架只能建立在广义相对论的基础上,其道理就在这里.

有趣的是,在以广义相对论为基础的宇宙动力学建立后,Milne 在 1934 年指出,设想宇宙是一个虽然很大但终究有限的均匀介质球,而我们的可观测宇宙处在非常远离边界的普通位置上,那么牛顿理论能提供出同样形式的方程式.当然,这样的图景与宇宙学原理是冲突的.有限的模型宇宙必有一个中心,不同位置也不可能平等.这就把宇宙哥白尼原理的思想破坏了.但是由于人们对牛顿力学比对广义相对论熟悉得多,所以下面用广义相对论讨论宇宙的动力学时,将指出它的牛顿类比.这对便于想像是有助的.无论如何,这两种理论在概念上是有重要区别的.

二、引力与时空的弯曲

要理解宇宙学,先对广义相对论有初步的了解是需要的(参看附录Ⅰ).让我们仍用电磁类比来了解.用今天理论物理的眼光看,把库仑的静电规律发展成麦克斯韦的一般电磁理论包含着两重内涵.首先是弄清,应当把由电势 ϕ 描写的静电场扩展为用电磁势 A^μ(即标势 ϕ 和矢势 A)描写的一般电磁场.然后是找出一般电磁势 A^μ 所满足的规律,即麦克斯韦方程组.当爱因斯坦要把牛顿的静引力规律发展成一般的引力理论,他需要做的是同样的两步.

无论如何,电磁理论和引力理论发展中条件的差别是很大的.在麦克斯韦建立电磁理论前,他已有电流(即电荷的运动)会产生磁场,以及磁场的变化会产生电场的经验规律作指引.因此今天看来,麦克斯韦理论的建立是一件水到渠成的事.而爱因斯坦要发展静引力理论时却完全不具有运动物体如何产生引力场的经验.于是他在很大程度上要依靠理性思维.

1907 年,爱因斯坦完成了第一步.他从引力质量和惯性质量等同的事实认识到,惯性力和引力是不可区分的.他由此断言,一

般的引力场应是牛顿引力和惯性力的统一描写，惯性力应是引力的一部分. 这就是他的等效原理. 作为等效原理的推论，一般引力场应当由四维时空度规 $g_{\mu\nu}$ 来描写. 从物理概念上讲，这结果意味着引力场的存在是通过时空的弯曲来体现的. 仅当引力场很弱，即时空弯曲很小，引力场才可以用平坦时空上叠加一个经典场来描述. 牛顿引力定律就是这种情形. 在一般情况下，引力场的存在将只体现为时空的弯曲. 这是爱因斯坦引力理论，即广义相对论的基本概念. 这一个概念不仅是人们对引力认识的深化，也是对时空认识的深化.

我们都已知道，狭义相对论的贡献是指出了时间与空间是一个整体. 但是狭义相对论依然把它们看成是平坦的、与物质世界无关的. 在这样的平坦时空中可取 $x^0=ct$ 代表时间，$x^1=x$、$x^2=y$ 和 $x^3=z$ 代表空间的直角坐标，则任两个时空坐标差为 $\mathrm{d}x^\mu$（$\mu=0,1,2,3$）的事件点之间的四维距离平方是我们熟知的勾股弦定理的推广，即

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2. \tag{2.1}$$

若对时空采用曲线坐标，则四维距离平方的一般形式将是

$$\mathrm{d}s^2 = \sum_{\mu,\nu} g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu. \tag{2.2}$$

其中的 $g_{\mu\nu}$ 就是上面提到的时空度规. (2.1)式相当于 $g_{00}=-1$，$g_{11}=g_{22}=g_{33}=1$，及当 $\mu\neq\nu$ 时有 $g_{\mu\nu}=0$. 这叫狭义相对论的 Minkowski 度规.

容易直观理解，在弯曲时空中是无法建立直角坐标系的. 因此，四维距离平方被一般地定义为(2.2)的形式，即时空的性质完全由度规 $g_{\mu\nu}$ 来确定. 这里引申出一个数学问题：那么如何区别弯曲时空和采用了曲线坐标的平坦时空？

时空的平坦或弯曲是一种客观属性. 它与坐标的选择应没有关系，因为坐标仅是人们研究时空性质的手段. 例如在研究二维平面时也可以采用曲线坐标，但这平面并不会因此而变成曲面. 按这

道理，我们可以用如下的判据来区别时空的弯曲或平坦：当能用时空坐标变换把(2.2)形式的距离平方还原成(2.1)的简单形式，则这度规所描述的时空是平坦的；如果办不到才说明时空是弯曲的.

从几何学讲，度规是描述空间(时空)是否弯曲以及如何弯曲的基本量.广义相对论则论证了时空的弯曲状况是引力场的体现.这样，度规成了描述引力场分布的物理量.引力场的分布当然是由物质决定的.因此按广义相对论的观念，时空不再是与物质世界无关的背景.反之，它的性质应当完全是由物质决定的.这在物理学中是一个崭新的概念.

弄清了应当用度规来描写引力场是建立广义相对论的第一步.同样要紧的第二步是找出引力场(即度规场)所满足的物理规律.如前所述，广义相对论前的物理规律在很大程度上都是来自实践经验的归纳.但是想找出一般引力场所适用的规律，却没有实践经验可借鉴.这使得爱因斯坦只能再靠理性的推测.

类似于一般电磁场的源是电荷的分布和运动一样，一般引力场的源应是物质的分布和运动.描写它的物理量是物质的四维能量——动量张量 $T_{\mu\nu}$. 现在需要的是找出引力场(即度规场 $g_{\mu\nu}$)对场源 $T_{\mu\nu}$的依赖关系.这关系应当表现为四维张量形式的偏微分方程.这里要达到的目标很具体，即找出引力场方程，但是却没有具体的经验作借鉴，这件事太困难了.爱因斯坦为此花了 8 年功夫.最后到 1915 年，在数学家 Grossmann 的帮助下才发现：如果在上面的考虑外，再要求场方程为二阶的偏微分方程，且对二阶微商项为线性，则所要的方程形式是惟一地确定的，即

$$R_{\mu\nu} - (1/2)Rg_{\mu\nu} - \lambda g_{\mu\nu} = - (8\pi G/c^4)T_{\mu\nu}. \qquad (2.3)$$

其中 $R_{\mu\nu}$和 R 是由度规 $g_{\mu\nu}$及其一阶和二阶微商所构成的、形式完全确定的张量.于是，爱因斯坦要找的场方程有了.

这样的方程中涉及三个物理常数.两个是已知常数，即牛顿引力常数 G 和光速 c. 另一个是任意常数 λ. 为使牛顿万有引力定律

能作为这个方程的好的近似，λ 必须是一个很小的常数，以致它在尺度不大(例如太阳系)的引力现象中不起作用. 爱因斯坦的做法是索性把 λ 取作零. 这样，引力场方程(2.3)就完全地确定了. 它就是现在广义相对论所用的爱因斯坦引力场方程.

这里必须注意，若 λ 不为零而是一个很小的常数，它对牛顿引力定律已得到证实的现象没有影响，而对宇宙尺度上的动力学影响将是不可忽视的. 由此人们把 λ 叫做宇宙常数. 它作为在基本物理方程中出现的常数，其大小只能从实测来确定，而原则上不可能由理论推断的. 因此，在研究宇宙动力学时，我们应当用 $\lambda=0$ 的爱因斯坦场方程，还是用保留宇宙常数 λ 的场方程，这是宇宙学观测研究的基本问题之一.

Friedmann 在 1923 年首先用 $\lambda=0$ 的爱因斯坦场方程建立了膨胀宇宙的模型. 后来多数研究者都沿用着它，因此这样的理论模型被称为宇宙学的标准模型. 我们在本章要讨论的就是这模型的性质及它的实践检验. 至 1998 年底，才有 Perlmutter 等人用天文观测初步断定 $\lambda>0$. 尽管这还不一定是最后的定论，无论如何，$\lambda\neq0$ 的模型引起了人们格外的注意. 在下一章中我们将专门讨论宇宙常数对宇宙动力学的影响. 要是 $\lambda>0$ 的结果将来得到观测的确认，那么以含宇宙常数 λ 的场方程为基础的模型将会成为新的标准模型.

三、均匀和各向同性的时空

不管从物理上还是从数学上讲，广义相对论都是一个相当复杂的理论. 但是我们的兴趣仅在于用它讨论宇宙的运动. 结果事情却是意外的简单.

我们已经知道，在宇观上，宇宙是一片充满全空间的均匀介质. 这样，由它产生的引力(度规)场也自然应当是均匀和各向同性的. 所谓均匀，指空间不同点上的度规应没有区别；而各向同性，则

指从一点度量不同方向应没有区别.正是这两方面要求,把度规的一般形式限定得很简单.

虽然我们在宇宙中的位置没有特殊性,但为研究方便,依然把坐标原点取在我们所在的银河系.其他质元(星系)的位置由广义的球坐标 r, θ 和 ϕ 标记.这里用的是随体坐标,意即在宇宙运动(膨胀或收缩)中,每一质元的坐标 r 是不变的.它与我们的距离的变化将由尺度因子 $R(t)$ 来描述.这样,纯用数学可证明,四维距离的平方总能写成

$$ds^2 = -c^2dt^2 + R(t)^2\left\{\frac{dr^2}{1-kr^2} + r^2d\theta^2 + r^2\sin^2\theta d\phi^2\right\}. \tag{2.4}$$

这种描写均匀各向同性时空的度规叫 Robertson-Walker 度规.当仅考虑到均匀和各向同性,式中的尺度因子 $R(t)$ 可以是任意函数.宇宙的具体运动情况应由动力学方程决定.此外式中还含有一个任意常数 k.其大小只能由实际测量来确定.

如果在平坦时空中采用空间球坐标,相应的四维距离平方为

$$ds^2 = -c^2dt^2 + dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\phi^2. \tag{2.5}$$

它与(2.1)式是等价的.与(2.4)式相比,差别就在于 $R(t)$ 和 k. $R(t)$ 的含义是宇宙空间在膨胀(或收缩).若 $R(t)$ 等于常数,则空间是静止的.下面让我们讨论一下 k 的含义.

先讨论 $k=0$ 的情形.这时若只看(2.4)的空间部分,它显然能够用坐标变换,把三维距离化成(2.5)或(2.1)中的相应形式.按前面的判据,这表明 $k=0$ 情形下的三维空间是平坦的.再看 $k\neq0$ 的情形.这时三维距离平方绝不可能化成(2.5)中的形式.同样的判据告诉我们,相应的三维空间是弯曲的.这样就定性地看清了,k 是反映宇宙三维空间是否弯曲的因子.值得顺便指出,不管 k 是否为零,这四维时空总是弯曲的.

定量地讲,空间的弯曲程度是由 $k/R(t)^2$ 描述的.由于 $R(t)$ 的大小是相对的,所以 k 本身的大小也没有确定的意义.因此我们只

须讨论 $k=+1,0$ 和 -1 三种情形就够了. 人们常简单地把它们叫做正曲率的、零曲率的(即平坦的)或负曲率的宇宙. 平坦空间的几何比较直观,下面对弯曲空间的性质作一些说明.

三维的弯曲空间很难直观. 让我们先讨论 $k=+1$ 的二维空间(即曲面). 在(2.11)式中去掉 ϕ 角,看出空间相邻两点的距离公式为

$$dl^2 = R(t)^2\left\{\frac{dr^2}{1-r^2} + r^2 d\theta^2\right\}. \tag{2.6}$$

注意这里坐标 r 的取值只能从 0 到 1. 作坐标变换

$$r = \sin\psi, \tag{2.7}$$

上述距离公式化成了

$$dl^2 = R(t)^2(d\psi^2 + \sin^2\psi d\theta^2). \tag{2.8}$$

由此看到,一个正曲率的二维空间相当于镶嵌在三维平坦空间中半径为 $R(t)$ 的二维球面. 借助二维的例子,我们直观地看到正曲率空间有几个有趣的特征:第一,对一个球面,显然一切点是平等的,从任一点看不同方向也是平等的,即符合宇宙学原理要求的. 第二,这空间的总体积(实为面积)是有限的,即 $4\pi R(t)^2$. 总质量因此也是有限的. 第三,这空间虽然是有限大,但却是没有边界的.

回到三维的正曲率空间. 用同样的变换可把空间间隔化为

$$dl^2 = R(t)^2\{d\psi^2 + \sin^2\psi d\theta^2 + \sin^2\psi\sin^2\theta d\phi^2\}. \tag{2.9}$$

这是镶嵌在四维平坦空间中半径为 $R(t)$ 的三维球"面". 由于四维平坦空间难以想像,其中的三维球"面"也失去了直观性. 但上面对二维正曲率空间的描述能帮我们理解,它同样具有上述三条性质. 特别强调,这宇宙空间总体积为 $2\pi^2R(t)^3$,即是有限的,但它却是没有边界的.

对于 $k=-1$ 的三维空间更难直观描述,我们不拟细细讨论. 主要注意一点差别,即这时径向坐标 r 的取值是 0 到∞,因此其总体积也是无穷大. 简言之,这是一个无限而弯曲的三维空间. 至于它的均匀和各向同性,则是在设计度规(2.4)时已保证了的.

总之，考虑到宇宙学原理的要求，宇宙的三维空间只有三类可能，即平坦的、负曲率的或正曲率的. 前两类是无限宇宙，而后一类则是有限宇宙. 这里讨论的是理论上的一切可能. 实际的宇宙只能是这三种可能类型中的一种. 至于它属于哪一种，这是必须由实测来回答的问题. 后面即将会讨论到.

四、膨胀的动力学

现在我们具体地研究宇宙在自身引力作用下的运动过程，从而把尺度因子 $R(t)$ 确定下来. 这首先要把爱因斯坦引力场方程具体化.

宇宙介质作为引力源，它被看成均匀的理想"气体". 其物理性质用密度 ρ 和压强 P 来描述. 在宇宙膨胀或收缩的过程中，随尺度因子 R 的变化，密度和压强也将在变化. 这就是说，ρ 和 P 也都是时间 t 的函数. 现在需要的是写出一组动力学微分方程，以作为求解 $R(t)$，$\rho(t)$ 和 $P(t)$ 的出发点. 考虑到所有有关的物理量都只与时间有关，而与空间点无关，所以所得到将注定是一组简单的常微分方程.

把宇宙的 Robertson-Walker 度规代入爱因斯坦场方程，在适当整理后，得出两个常微分方程如下：

$$\dot{R}^2 + k = \frac{8\pi G}{3}\rho R^2, \tag{2.10}$$

$$\ddot{R} = -\frac{4\pi G}{3}(\rho + 3P)R. \tag{2.11}$$

式中的 G 是牛顿引力常数，变量上的附点代表对时间的微商. 注意从这里开始，我们将一直采用 $c=k=\hbar=1$ 的自然单位制①，因

---

① 参看附录Ⅰ. 这里只须先注意到一点. 由于令光速 $c=1$，质能关系简化为 $E=m$，因此质量和能量是不必区分的. 此外，动能、动量、能流、动量流或压强也都与质量有相同的量纲.

作为基本的动力学方程. 它被称为 Friedmann 方程. 这方程涉及 $R(t)$和 $\rho(t)$. 在实物为主的前提下,辅以 ρR^3=常量,方程组就完备了.

为了要得到确定的解,还须配以适当的初条件. 通常定义膨胀从 $R=0$ 开始,并把相应的时刻记作 $t=0$. 这样,初状态是 $\rho=\infty$和 $R=0$. 它们不构成一组适合求解的初条件. 我们知道,为使力学问题有惟一解,所用的初条件不一定与起始状态相对应. 它完全可以是过程中任何一个时刻的状态. 从宇宙学讲,较容易用测量确定的是今天的状态. 所以下面将以今天为"初"来求过去的运动. 这样的初条件实际上是"终"条件,但对于求解是同样有效的.

我们将对任何物理量加下标 0 代表它在今天的值,如今天的尺度因子为 R_0, 膨胀速率为 H_0,物质密度为 ρ_0 等. 利用

$$\rho R^3 = \rho_0 R_0^3, \tag{2.16}$$

Friedmann 方程可重写作

$$\left(\frac{\dot{R}}{R_0}\right)^2 + \frac{k}{R_0^2} = \frac{8\pi G}{3}\frac{\rho_0 R_0}{R}. \tag{2.17}$$

这是尺度因子 $R(t)$满足的微分方程. 注意到今天尺度因子 R_0 的大小是可以任意规定的,所以需要求解的是 $R(t)/R_0$. 方程中实际只含有两个常参量 ρ_0 和 k,需要用实测值输入. 把上述方程式用于今天,可得到一个代数关系,

$$H_0^2 + \frac{k}{R_0^2} = \frac{8\pi G}{3}\rho_0, \tag{2.18}$$

其中 $H_0=\dot{R}_0/R_0$,它就是哈勃常数. 从这关系看出,三个常数 k, H_0 和 ρ_0 来中只有两个是独立的. 我们可以根据需要来选用.

让我们用 k 和 H_0 作输入参量. 在消去 ρ_0 后,可以容易地把 Friedmann 方程解出:

$$t = \frac{1}{H_0}\int_0^{R(t)/R_0} \sqrt{\frac{x}{1+\frac{k}{R_0^2H_0^2}(1-x)}}\,\mathrm{d}x. \tag{2.19}$$

作为基本的动力学方程. 它被称为 Friedmann 方程. 这方程涉及 $R(t)$和 $\rho(t)$. 在实物为主的前提下,辅以 $\rho R^3=$常量,方程组就完备了.

为了要得到确定的解,还须配以适当的初条件. 通常定义膨胀从 $R=0$ 开始,并把相应的时刻记作 $t=0$. 这样,初状态是 $\rho=\infty$和 $R=0$. 它们不构成一组适合求解的初条件. 我们知道,为使力学问题有惟一解,所用的初条件不一定与起始状态相对应. 它完全可以是过程中任何一个时刻的状态. 从宇宙学讲,较容易用测量确定的是今天的状态. 所以下面将以今天为"初"来求过去的运动. 这样的初条件实际上是"终"条件,但对于求解是同样有效的.

我们将对任何物理量加下标 0 代表它在今天的值,如今天的尺度因子为 R_0, 膨胀速率为 H_0,物质密度为 ρ_0 等. 利用

$$\rho R^3 = \rho_0 R_0^3, \tag{2.16}$$

Friedmann 方程可重写作

$$\left(\frac{\dot{R}}{R_0}\right)^2 + \frac{k}{R_0^2} = \frac{8\pi G}{3}\frac{\rho_0 R_0}{R}. \tag{2.17}$$

这是尺度因子 $R(t)$满足的微分方程. 注意到今天尺度因子 R_0 的大小是可以任意规定的,所以需要求解的是 $R(t)/R_0$. 方程中实际只含有两个常参量 ρ_0 和 k,需要用实测值输入. 把上述方程式用于今天,可得到一个代数关系,

$$H_0^2 + \frac{k}{R_0^2} = \frac{8\pi G}{3}\rho_0, \tag{2.18}$$

其中 $H_0=\dot{R}_0/R_0$,它就是哈勃常数. 从这关系看出,三个常数 k, H_0 和 ρ_0 来中只有两个是独立的. 我们可以根据需要来选用.

让我们用 k 和 H_0 作输入参量. 在消去 ρ_0 后,可以容易地把 Friedmann 方程解出:

$$t = \frac{1}{H_0}\int_0^{R(t)/R_0} \sqrt{\frac{x}{1+\frac{k}{R_0^2H_0^2}(1-x)}}\,\mathrm{d}x. \tag{2.19}$$

作为基本的动力学方程. 它被称为 Friedmann 方程. 这方程涉及 $R(t)$和 $\rho(t)$. 在实物为主的前提下，辅以 $\rho R^3=$常量，方程组就完备了.

为了要得到确定的解，还须配以适当的初条件. 通常定义膨胀从 $R=0$ 开始，并把相应的时刻记作 $t=0$. 这样，初状态是 $\rho=\infty$和 $R=0$. 它们不构成一组适合求解的初条件. 我们知道，为使力学问题有惟一解，所用的初条件不一定与起始状态相对应. 它完全可以是过程中任何一个时刻的状态. 从宇宙学讲，较容易用测量确定的是今天的状态. 所以下面将以今天为"初"来求过去的运动. 这样的初条件实际上是"终"条件，但对于求解是同样有效的.

我们将对任何物理量加下标 0 代表它在今天的值，如今天的尺度因子为 R_0，膨胀速率为 H_0，物质密度为 ρ_0 等. 利用

$$\rho R^3 = \rho_0 R_0^3, \tag{2.16}$$

Friedmann 方程可重写作

$$\left(\frac{\dot{R}}{R_0}\right)^2 + \frac{k}{R_0^2} = \frac{8\pi G}{3}\frac{\rho_0 R_0}{R}. \tag{2.17}$$

这是尺度因子 $R(t)$满足的微分方程. 注意到今天尺度因子 R_0 的大小是可以任意规定的，所以需要求解的是 $R(t)/R_0$. 方程中实际只含有两个常参量 ρ_0 和 k，需要用实测值输入. 把上述方程式用于今天，可得到一个代数关系，

$$H_0^2 + \frac{k}{R_0^2} = \frac{8\pi G}{3}\rho_0, \tag{2.18}$$

其中 $H_0=\dot{R}_0/R_0$，它就是哈勃常数. 从这关系看出，三个常数 k，H_0 和 ρ_0 来中只有两个是独立的. 我们可以根据需要来选用.

让我们用 k 和 H_0 作输入参量. 在消去 ρ_0 后，可以容易地把 Friedmann 方程解出：

$$t = \frac{1}{H_0}\int_0^{R(t)/R_0} \sqrt{\frac{x}{1+\frac{k}{R_0^2 H_0^2}(1-x)}}\,\mathrm{d}x. \tag{2.19}$$

得到的是所要的 $R=R(t)$ 的反函数形式. 它描写了宇宙的膨胀过程.

图 2.2 按 k 值的不同,画出了不同的膨胀图景. 当 k 为正,膨胀阶段是有限的. 到一定时刻,宇宙将转向收缩,直至回到 $R=0$. 当 k 等于或小于零,尺度因子永远单调地增大. 整个膨胀过程是无限的. 宇宙密度一直在降低,直至介质变得无限稀薄.

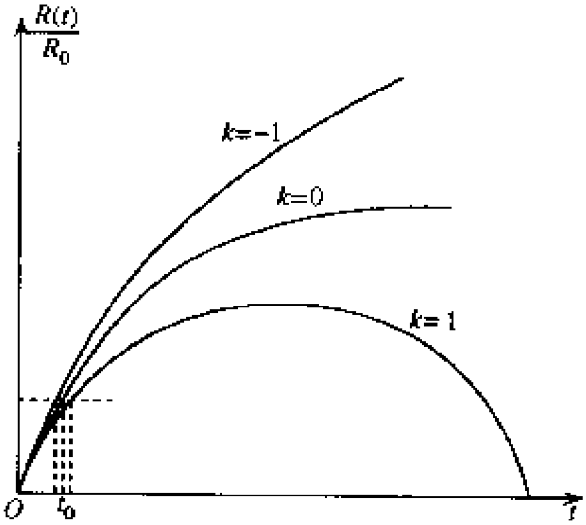


图 2.2 宇宙的膨胀过程

如前面提到过,这是从广义相对论得到的结果,但是也可以用牛顿力学中的机械能守恒来理解. 记得已指出过: 质元的总机械能 E 与曲率因子 k 是反号的. 宇宙的膨胀相当于任一质元被逆着引力方向外抛. 当质元以负的总能逆着引力方向运动,它必定会转向回落. 这是机械能守恒的后果. 它与 k 为正的情况下的图景是一致的. 当质元以正的总能逆引力方向而动,同样按机械能的守恒性,它必无限地远离. 这结果正就是 k 为零或正的情景. 这类比能使我们对上面的结果理解得直观一些.

在确切地解出 $R(t)$ 后,过去任一时刻的 R 大小(相对于 R_0)就知道了. 这在讨论许多宇宙学问题时是很关键的. 无论如何,使用像(2.19)那样的解式不很方便. 在 $k=0$ 的情况下,这积分容易解析地积出. 它是

$$R(t)/R_0 = (3H_0t/2)^{2/3}. \tag{2.20}$$

这是一个很简单的关系：时间增大三个量级，则尺度因子增大两个量级. 从图上的曲线看，当只讨论宇宙过去的状况，k 不同造成的差别并不大. 因此在仅需要作粗略估算时，不管 k 实际多大，用这个解是方便而合理的近似. 我们将会用到它.

最后让我们讨论宇宙膨胀速率和加速率随时间的变化.

任意 t 时刻的膨胀速率用

$$H(t) = \dot{R}(t)/R(t) \tag{2.21}$$

来描写. 它在今天的值就是哈勃常数. 由于引力的影响，过去的速率应比今天的小. 以 $k=0$ 为例看，由(2.20)式算出

$$H(t) = 2/3t. \tag{2.22}$$

膨胀速率是与时间成反比地降低的. 特别是，$t=0$ 时的膨胀速率是无穷大. 由于引力的减速，才使得今天的速率变成了 H_0.

再进一步看宇宙膨胀的加速度. 人们常用一个无量纲量

$$q = -\ddot{R}R/\dot{R}^2. \tag{2.23}$$

来描写加速度. 定义中添了个负号，使减速情况下的 q 值是正的. 这仅是为了方便. 它因此也被称为减速因子. 在方程(2.11)中代入 $P=0$(即实物为主)，易于由此导出

$$q = 4\pi G\rho/3H^2. \tag{2.24}$$

这式子说明，任一时刻的加速度 q 可用当时的密度 ρ 和速率 H 推出. 原则上，今天的加速度 q_0 是可以测量的. 实际上做这种测定很困难，因此我们对它的大小知道得很粗略. 例如能看到这样的结果：$q_0=1.0\pm0.8$. 反正今天的 q_0、H_0 和 ρ_0 中只有两个是独立的. 测量了任两个，就能推知第三个. 如三个都有了分别的测定，就能用它们来检验模型的正确性.

最后强调一点. 本节讨论的是实物为主的宇宙解. 以后将会论证，宇宙演化的早期肯定是以辐射为主的. 因此这里的解对它是不适用的. 辐射为主的早期约持续了 10^4 年. 因此在有的问题的讨论

中，这短暂的阶段是可以忽略的.

六、有限或无限

前面已讨论过，按照广义相对论，宇宙空间可能是无限的，也可能是有限的. 现在让我们考虑如何用测量来判定实际宇宙是无限或有限的问题. 要在宇宙尺度上作几何测量是目前办不到的事. 但是曲率因子 k 为正，或为负(包括零)却是能用其他物理量的测量结果来推定的.

利用哈勃常数 H_0 是可观测量，定义一个密度量纲的量

$$\rho_c = \frac{3H_0^2}{8\pi G}. \tag{2.25}$$

它也是能用观测确定的量(习惯上不加下标 0). 把 H_0 写作

$$H_0 = 100h \text{ km} \cdot \text{s}^{-1} \cdot \text{Mpc}^{-1}. \tag{2.26}$$

H_0 的不确定性表现为

$$h = 0.5\text{—}0.8. \tag{2.27}$$

它实际上是以 100 km · s^{-1} · Mpc^{-1}为单位的无量纲哈勃常数. 这样，ρ_c 的大小为

$$\rho_c = 1.9h^2 \cdot 10^{-29} \text{ g/cm}^3. \tag{2.28}$$

粗略说来，相当于每立方米有几个氢原子的密度.

引入了 ρ_c，由 Friedmann 方程导出的(2.18)式可写成

$$\frac{3k}{8\pi G R_0^2} = \rho_0 - \rho_c \tag{2.29}$$

从这关系式看出，若今天的宇宙密度比 ρ_c 大，则表明曲率因子 k 是正值，即宇宙空间的总容积是有限的；反之若实际密度等于 ρ_c 或比 ρ_c 小，则表明 k 是零或负数，即宇宙空间的总容积是无穷. 这样我们就有了一个回答宇宙为有限或无限的判据. ρ_c 是这判据中的临界量，因而它被称为临界密度. 宇宙学家常引入无量纲密度

$$\Omega_0 = \rho_0/\rho_c. \tag{2.30}$$

来代替 ρ_0. 这样, $\Omega_0=1$ 就是宇宙有限或无限的临界值.

要使用这判据,须先测定哈勃常数 H_0 和宇宙密度 ρ_0. 可是实际上我们对这两个量都知道得不很确切. H_0 的实测范围已写在上面了. 它使 ρ_c 的大小有超过一倍的不确定性. 人们对宇宙密度了解得更差. 天文学家已清楚地知道,宇宙中绝大部分物质并不发光. 用能较准确地作测量的方法时,很可能大部分物质没有被测进去. 用肯定能测到全部物质的方法时,测定的可靠程度常很低. 这问题将在暗物质一章中专门讨论. 今天肯定知道的是一个范围:一头是 ρ_0 肯定大于 $0.2\rho_c$,另一头是不排除 ρ_0 比 ρ_c 略大的可能.

把这样的测量结果与上述判据对照,宇宙空间为无限或有限的问题尚不能有明确的答案. 或许在今后的一二十年里,天文学家将会给出回答. 无论如何,上面的分析清楚地说明,宇宙的有限或无限是一个观测宇宙学应当能够,且肯定能够回答的具体问题,而不是哲学问题. 要说哲学家能可靠地告诉我们答案,无异于说哲学家能不经测量而知道今天的宇宙密度有多大. 这实在是不可思议的事.

上面讨论的只是一种判定方法. 其实理论上还有其他方法. 例如把微分方程(2.11)用于实物为主的今天,得出宇宙减速因子 q_0 与无量纲宇宙密度 Ω_0 有关系

$$q_0 = \Omega_0/2. \tag{2.31}$$

这样 $q_0=1/2$ 是临界情形. 若它实际上大于 1/2,则宇宙是有限的;反之则是无限的. 如已指出过,今天对 q_0 的测定值更不准,因此不详细讨论了.

七、宇宙的年龄

宇宙总体积可能有限常使人感到难以想像,而宇宙有有限年龄则曾使人们感到难以接受. 但是它是宇宙学理论的后果而不是前提. 从前面解出的膨胀过程看,由今天往前追溯是有尽头的. 它

就是 $R=0$，即密度为无穷的状态．把这时刻规定为时间的起点，那么今天宇宙的年龄为有限大是不可避免的推论．让我们从理论上算出这年龄的大小，并讨论用实测来检验这结果的办法．

第五节中已解出了描述膨胀过程的 $R(t)$ 的函数形式．令 $R(t)=R_0$，相应的时间 t_0 就是今天的宇宙年龄．在 $k=0$（即平坦宇宙）的情况下，$R(t)$ 的解式最简单．由（2.20）式立即看到

$$t_0 = (2/3)H_0^{-1}. \tag{2.32}$$

当然今天人们尚不肯定宇宙平坦与否，因此需要的是一般结果．它可以写成

$$t_0 = H_0^{-1}F(\Omega_0), \tag{2.33}$$

其中 Ω_0 是今天宇宙的无量纲密度．$F(\Omega_0)$ 对 Ω_0 的依赖关系画在图 2.3 中．注意它是 Ω_0 的降函数．$\Omega_0=1$ 代表平坦宇宙，图上相应读出的值是 $F(\Omega_0=1)=0.67$，即 2/3．由于 $F(\Omega_0)$ 总小于 1，所以 H_0^{-1} 是宇宙年龄的一个可靠的上限．

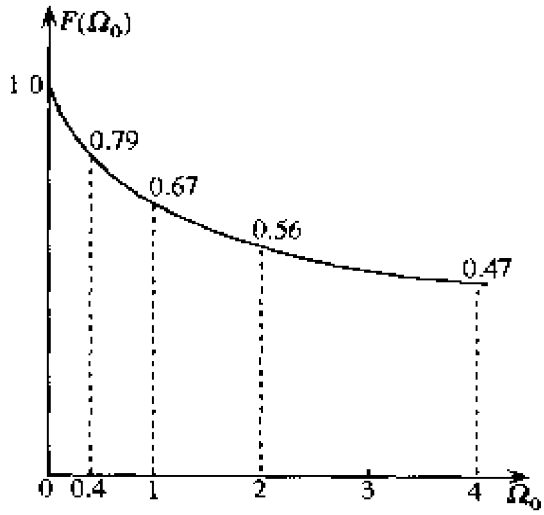


图 2.3　函数 $F(\Omega_0)$

从这结果看，当我们有 H_0 和 Ω_0 的实测值，那么就能推断出理论上的宇宙年龄了．可是如前面已提到，今天对这两个值都知道得不确切，所以只能定出一个较宽的范围：

$$t_0 = 10\text{—}20\ \text{Ga}, \tag{2.34}$$

即 100 至 200 亿年. 若最终确定下来的 H_0 和/或 Ω_0 值偏大，则理论的宇宙年龄偏向上述范围中小的一侧.

现在转向宇宙年龄如何检验的问题.

宇宙年龄 t_0 本身没有直接测量的办法，但是天文学家能测量古老天体的年龄 τ_0. 若该天体形成于宇宙年龄为 t_1 时，则它们有关系

$$t_0 = t_1 + \tau_0. \tag{2.35}$$

对于足够古老的天体，有 $t_1 \ll \tau_0$. 因此它的年龄 τ_0 可作为宇宙年龄 t_0 的近似. 这就是人们测定宇宙年龄的方法.

球状星团是天文学家能较准确地测定其年龄的最古老的天体. 既有方法上的原因也有测定上的技术困难，使得对不同球状星团测到的年龄很弥散. 它是

$$\tau_0 = 12—18\ \mathrm{Ga}, \tag{2.36}$$

即 120 至 180 亿年. 它与理论年龄是相洽的. 值得注意的是，在测量的精度提高后，两者出现矛盾的可能也是存在的. 例如若最终确定宇宙是平坦的($\Omega_0=1$)，而 H_0 的值是 70 $\mathrm{km \cdot s^{-1} \cdot Mpc^{-1}}$，那么理论的宇宙年龄将是 93 亿年. 它与球状星团的年龄肯定冲突了. 这问题在下一章谈含宇宙常数的模型时再讨论.

最后补充一个历史情况. 1952 年前，哈勃由于定距离的方法上的错误，给出的 H_0 的值是 500 $\mathrm{km \cdot s^{-1} \cdot Mpc^{-1}}$，即 H_0^{-1} 为 20 亿年. 按上面的分析，宇宙年龄必定小于 H_0^{-1}. 当时已经知道，太阳作为并不算古老的恒星，它的年龄为 50 亿年. 这与理论的宇宙年龄显然地冲突了. 值得指出的是，宇宙学原理作为理论模型的基本出发点，它在当时尚没有事实证据. 因此即使不管人们有拒绝接受宇宙有有限年龄的心态，对一个前提没有证据而推论又不符合事实的理论，它被学术界排斥是自然的事. 至 50 年代末，天文测量改正了哈勃的错误，肯定了 H_0 值是在 50 至 100 $\mathrm{km \cdot s^{-1} \cdot Mpc^{-1}}$之间，即宇宙年龄的上限在 100 至 200 亿年之间. 这样暂时就没有明显的年龄矛盾了. 但是这矛盾的缓解并不足以使宇宙学

理论赢得学术界的信任. 事实上是10年后背景辐射的发现,才成了宇宙学发展史上的转折点.

八、红移与距离的关系

第一章中已作为基本事实指出,远处星系光谱的红移量 z 与它和我们的距离 d_L 成正比. 这就是哈勃定律. 现在我们讨论相应的理论结果是什么. 定性地讲,光谱的红移来自宇宙的膨胀,它使星系在向远离我们的方向运动. 但是,定量的结果却与(非相对论的或相对论的)多普勒效应并不一样. 这是宇宙理论以广义相对论为基础的后果. 因此红移与距离的关系是检验宇宙学模型的一个重要方面.

首先讨论一下距离. 天文学家用光度推断距离的依据是(参看(1.1))

$$B = \frac{L}{4\pi d_L^2}. \tag{2.37}$$

这里对距离 d 加下标 L,以表明是由光度 L 推断的. 关系式(2.37)中暗含了一个很普通的假定: 空间是平坦的,服从欧氏几何的. 而从广义相对论看,这假定却是不成立的. 当空间为弯曲,(2.37)中的 d_L 并不是几何距离. 但是天文学家实际上是由这关系式来告诉我们远处星系的"距离"的. 因此让我们在名称上区分一下,把这样得到的结果叫星系的光度距离. 千万记得: 天文学家测到的不是几何距离. 理论上可证明,光度距离与宇宙学量的关系是

$$d_L = R_0 r_1 (1 + z), \tag{2.38}$$

其中 r_1 是远处星系的径向坐标(我们自身的径向坐标为0),z 是星系的红移.

然后讨论红移. 星系光谱的红移 z 被定义为

$$z = \frac{\nu_0 - \nu}{\nu}, \tag{2.39}$$

其中 ν_0 是光被发射时的频率，ν 是接收者测到的频率. 由于光的传播需要时间，因此我们在 t_0 时接受到的光是星系在 t_1 时发出的. 当然 t_1 必比 t_0 小. 研究光在膨胀宇宙中的传播，可以得到

$$1 + z = R_0/R(t_1), \tag{2.40}$$

此外，标志星系位置的坐标 r_1 和发光时间 t_1 有明确的依赖关系. r_1 越大代表星系越远，则 t_1 须越小，才能使我们在 t_0 时(即今天)接受到它的光. (2.40)式又告诉我们，测到的红移 z 也越大. 由于星系的坐标 r_1 不是可观测量，因此我们可用星系红移 z 的大小作为远近的标志. 原则上，r_1 和 t_1 都是被 z 决定的. 具体地导出其关系，代入(2.38)式就可得到光度距离 d_L 与红移 z 的关系. 这两个量都能够实测，因此理论结果可以检验.

这里我们对于推导细节不感兴趣，直接写出结果如下：

$$H_0 d_L = \frac{2}{\Omega_0^2}\left[\Omega_0 z - (2 - \Omega_0)(\sqrt{\Omega_0 z + 1} - 1)\right] \tag{2.41}$$

它比经验的哈勃定律要复杂得多. 当对红移较小的星系，可把上述结果展开成 z 的幂级数，

$$H_0 d_L = z + \frac{1}{2}\left(1 - \frac{\Omega_0}{2}\right)z^2 + \cdots \tag{2.42}$$

哈勃定律就是从第二项起都可忽略时的结果①. 让我们看如何检验它.

我们已经知道，要有确定的、可检验的理论结果，需要先输入两个“初”条件. 通常被取用的是 H_0 和 Ω_0(代替 ρ_0). 现在看到，要把距离-红移关系式(2.41)完全确定，同样需要这两个值. 这是意料之中的. 因此检验的第一步是证实 z 与 d_L 有正比关系，并从而把 H_0 定下来. 这一步已通过对大量小红移星系的测量而确证到了无可置疑的程度. H_0 值的不确定性主要是距离测量中的系统误差带来的，至今尚无法排除. 进一步的检验是对红移不很小的星系

① 这理论公式中取用了光速 $c=1$. 按普通单位制，应为 $H_0 d_L = cz$.

作测量，看光度距离与红移的（非线性）关系是否符合(2.41)的形式. 为理解做这一步检验的困难，让我们估算一下有关星系该有多远.

对红移 z 小于 0.2 的星系，实测已证明哈勃的线性关系是成立的. 取 H_0 为 60 $\mathrm{km \cdot s^{-1} \cdot Mpc^{-1}}$ 来估算，$z=0.2$ 的星系的光度距离为 1000 Mpc (300 亿光年). 宇宙学中常把红移显著地小于 1 的星系叫“不太远”的星系. 从这里看到，它们的客观距离可能是相当远的. $z=1$ 的星系的光度距离约达 5000 Mpc，而距离-红移关系偏离线性的程度尚仅为 20%左右. 因此这样的检验要求对非常远的星系的光度距离测得很准，例如误差小于 10%. 这正是困难的所在. 无论如何，现在已有人初步地做到了这一点. 因为所得到的结果表明宇宙常数不为零，所以我们在下一章中再讨论它.

第三章　含宇宙常数的模型

> 长期来，多数宇宙学理论研究采用宇宙常数为零的标准模型.这主要是为了少引入不确定因素.当天文观测给出了宇宙常数不为零的证据，那么理论家只能修改标准模型，把宇宙常数的影响考虑进去，并重新来检验它.

一、宇宙常数的影响

对宇宙的动力学研究是以广义相对论为基础的.长期来，多数研究者采用宇宙常数 $\lambda=0$ 的爱因斯坦引力场方程.这样的模型被称为标准宇宙模型.叫它“标准”模型，只因为大家都这么做，而并不表明它是惟一正确的模型.近两年来，已有实测初步推定了非零的宇宙常数值.此外，宇宙年龄和结构形成的研究也显示出需要非零宇宙常数的迹象.所以，以含宇宙常数项的场方程为出发点的模型很可能会逐渐被大家所接受，而变成将来的标准模型.为此，我们有必要讨论一下宇宙常数对宇宙行为的影响.

当做这样的讨论时，宇宙时空用 Robertson-Walker 度规描写这点没有改变.受影响的仅是 $R(t)$ 所满足的动力学方程.现在须把含宇宙常数的项 $\lambda g_{\mu\nu}$ 加进去.这样，原来的方程(2.10)和(2.11)相应地改成了

$$\dot{R}^2 + k = \frac{8\pi G}{3}(\rho_m + \rho_{eff})R^2, \tag{3.1}$$

$$\ddot{R}^2 = -\frac{4\pi G}{3}(\rho_m - 2\rho_{eff})R, \tag{3.2}$$

其中的

$$\rho_{\text{eff}} = \lambda/8\pi G \tag{3.3}$$

具有密度的量纲. 它是宇宙常数的代表,也常被叫做等效真空能密度. 这名称的含义将在第九章中讨论. 这里先看方程中的附加项的物理意义.

先看修改后的 Friedmann 方程(3.1). 等式右边前一项 ρ_m 代替了原来的 ρ,仍表示实物的密度. 它满足 $\rho_m R^3$=常数. 差别在于现在在方程右边新加了一项膨胀中保持不变的密度 ρ_{eff}. 这项可理解成真空对位能的附加贡献. 然后看方程(3.2). 记得这方程相当于牛顿第二定律. 等式右边第一项是普通物质产生的引力. 由于实物为主,压强的贡献已略去了. 宇宙常数的影响体现在附加的第二项上. 值得注意,等效真空能产生的是斥力. 实物对宇宙膨胀起减速作用,而它起的是加速作用.

在上一章讨论宇宙常数为零的标准模型时,把 Friedmann 方程化成(2.17)的形式. 现在可由(3.1)化出相应的方程,其形式是

$$\left(\frac{\dot{R}}{R_0}\right)^2 + \frac{k}{R_0^2} = \frac{8\pi G}{3}\left(\frac{\rho_{m_0} R^2}{R_0^2} + \frac{\rho_{\text{eff}} R^2}{R_0^2}\right). \tag{3.4}$$

从这方程能解出 $R(t)$. 它描写了含宇宙常数模型中的膨胀过程.

原来解出的 $R(t)$ 依赖于两个输入参量. 现在新涉及的宇宙常数也是一个待定的参量. 这样,含宇宙常数模型中的解依赖三个输入参量. 通常被取用的是 H_0, Ω_{m_0} 和 Ω_λ. Ω_{m_0} 和 Ω_λ 分别是以临界密度为单位的实物密度和等效真空能密度,即

$$\Omega_{m_0} = \frac{\rho_{m_0}}{\rho_c} = \frac{8\pi G \rho_{m_0}}{3H_0^2}, \tag{3.5}$$

$$\Omega_\lambda = \frac{\rho_{\text{eff}}}{\rho_c} = \frac{\lambda}{3H_0^2}. \tag{3.6}$$

后者代替的是宇宙常数 λ. 当取用这组参量,相应的运动解为

$$t = \frac{1}{H_0}\int_0^{R(t)/R_0} \frac{dx}{\sqrt{1 + \Omega_{m_0}\left(\frac{1}{x} - 1\right) + \Omega_\lambda(x^2 - 1)}} \tag{3.7}$$

它仍然是 $R=R(t)$ 的反函数形式.

从形式上看,含宇宙常数的模型与标准模型的差别很小,因此它只是标准模型的一个简单的变种.实质问题在于:宇宙常数项的出现对宇宙膨胀过程是否有重大的影响?显然答案取决于 Ω_λ 的数值大小.

爱因斯坦最初就认识到:牛顿引力定律在太阳系内精确地成立,意味着等效真空能密度 ρ_{eff} 必定比太阳系的平均密度($\sim10^{-6}$ g/m^3)要小很多.但是会对宇宙膨胀起影响的密度是 ρ_c 的量级($\sim10^{-24}$ g/m^3).由此看来,真空能在宇宙动力学中也可能会起很重要的作用.如若 ρ_{eff} 比 ρ_c 大很多,那么不含宇宙常数的标准模型将显著地与事实不符.而我们在上一章中已看到,标准模型的行为大体与事实是相接近的.这表明 ρ_{eff} 至多只会与 ρ_c 同量级.实际上十几年前的宇宙学测量已给出了更强的限制:

$$\Omega_\lambda \leqslant 0.9. \tag{3.8}$$

这就是今天讲宇宙常数"很小"的具体含义.

从另一方面看,若 ρ_{eff} 显著地小于 ρ_c,即 $\Omega_\lambda \ll 1$,那么它对宇宙学的影响就不值得考虑了.让我们用 $\Omega_\lambda=\Omega_{m_0}<1$ 的例子来讨论.从基本方程(3.1)看出,今天真空能与实物同等重要.但是实物能密度反比于 R^3 减小,而真空能密度却是不变的.这样追溯到 $R(t)=R_0/2$ 时,真空能已比实物能约小一个量级.在更早的时候,这真空能将完全不起作用.因此对宇宙整个演化史,Ω_λ 的出现只对标准模型起了一个很小的修正作用.这分析说明,要在宇宙学理论中值得考虑 Ω_λ 的存在,其大小须满足

$$\Omega_\lambda \gtrsim 0.2. \tag{3.9}$$

简而言之,若宇宙常数 λ 的值使 Ω_λ 接近(3.9)所规定的下限或更小,它对宇宙演化的影响是很微小的;若它接近由(3.8)式规定的上限,则它将对标准模型产生可观测到的修正.宇宙常数在宇宙学理论中并没有很大的游移余地.

二、平坦性的判据

在标准模型中我们已经知道：实物密度 $\Omega_0=1$ 反映宇宙空间是平坦的；若它小于 1，空间是弯曲而无限的；若它大于 1，则空间是弯曲而有限的. 在含宇宙常数模型里，这判据只需要很小的修改.

把动力学方程(3.4)用于今天，得到的代数关系是

$$\frac{k}{R_0^2H_0^2}=\Omega_{m_0}+\Omega_\lambda-1. \tag{3.10}$$

令

$$\Omega_0=\Omega_{m_0}+\Omega_\lambda, \tag{3.11}$$

则得到了形式上与标准模型同样的判据：$\Omega_0=1$ 反映 $k=0$，即宇宙空间是平坦的. 若 Ω_0 大于或小于 1，其后果也与标准模型的预言一样.

这结果使我们可以把真空看成宇宙介质的一部分，从而把 Ω_0 理解为宇宙的总密度，即实物与真空的密度之和. 这样，标准模型用总密度判断宇宙空间为有限或无限的论据依然成立.

顺便讨论一下含宇宙常数模型中宇宙膨胀的减速因子 q_0 与密度的关系.

减速因子的定义见(2.23)式. 把方程(3.2)用于今天，得到一个代数关系

$$q_0=\frac{1}{2}\Omega_{m_0}-\Omega_\lambda. \tag{3.12}$$

实物密度越大，则减速因子越大. 这是很好理解的引力效果. 宇宙常数产生斥力，它对膨胀的效果是加速，式中的负号反映了这一点. Ω_λ 越大，则 q_0 越小. 上节中谈到，若宇宙常数的效应能被观测到，应有 $\Omega_\lambda \gtrsim \Omega_{m_0}$. 由(3.12)式看出，这样 q_0 将是负数，即今天宇宙的膨胀是加速的. 由于减速因子不直接与总密度相联系，所以从

q_0 的大小不再能判断空间的有限或无限. 但是若定出 q_0 为负数，则这是宇宙常数不为零的强烈证据，因为实物不可能使宇宙加速膨胀.

三、真空能密度的测定

现在继续讨论上一章末尾没有讨论完的红移-距离关系问题. 我们已经知道，困难在于可靠地测量上千 Mpc 的星系距离. 此外也知道，在检验红移-距离关系的同时，能够把宇宙介质的密度定出来.

90 年代初，人们开始发现，用 Ia 型超新星(记作 SNIa)做星系的距离指示器，能较可靠地测量几百 Mpc 量级的距离. 用这方法的不利之处是星系中出现 SNIa 的机会较小. 于是有了几个天文研究组专门搜索远处星系中的超新星，并得到了一些新结果. 例如在 1996 年，有一组人用 $z<0.1$ 的星系的红移-距离关系，定出了(图 3.1)

$$H_0 = 64 \pm 3 \text{ km} \cdot \text{s} \cdot \text{Mpc}^{-1}. \tag{3.13}$$

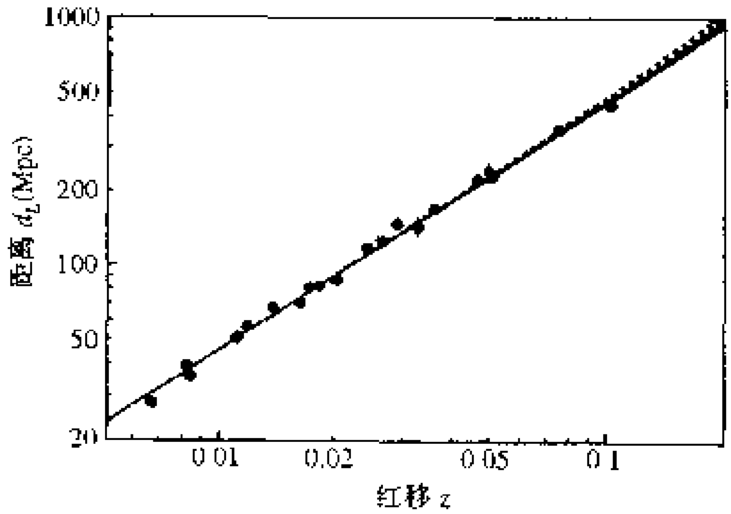


图 3.1　红移与距离的线性关系($z\leqslant 0.1$)

由于这种测定方法的系统误差依然不清楚，所以这还不是最终的结果. 但是专家们相信，这样的结果与哈勃常数的实际值已不会差太多了.

重要的进展出现在 1998 年末. Perlmutter 等人用 SNIa 方法积累了几十个较高红移($z>0.4$)星系的距离测量结果，使得由红移-距离关系较可靠地推断宇宙密度有了可能. 图 3.2 画出的是 Perlmutter 等人的测量结果和相应的理论拟合. 当用宇宙常数为零的标准模型公式(2.41)拟合时发现，所定出的密度 Ω_0 是负数(或很接近于零). 它显然与事实不符，从而强烈地暗示了宇宙常数不为零. 因此，这样的测量结果需要用含宇宙常数的模型来讨论. 相应的理论结果是早已得到了的.

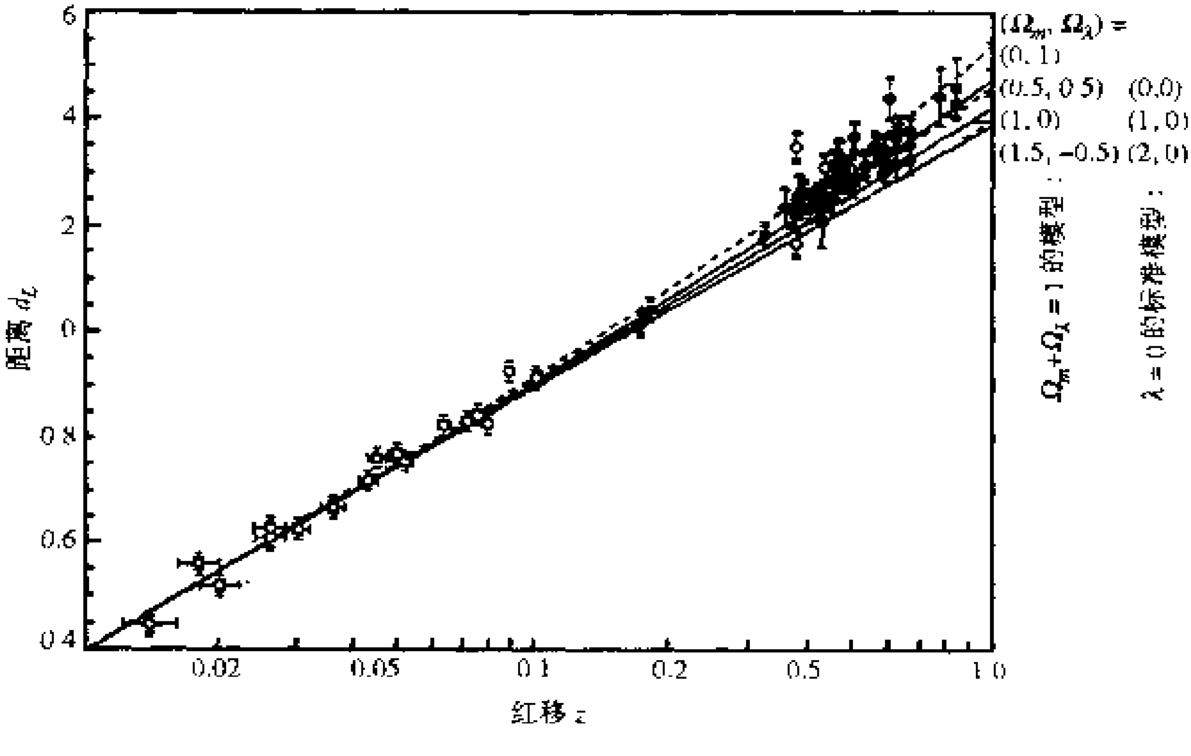


图 3.2 用 SNIa 定出的红移-距离关系

在含宇宙常数的模型中，理论上的红移-距离关系完全可以用同样的方法来导出. 所得到的关系有形式

$$H_0 d_L = F(\Omega_{m_0}, \Omega_\lambda). \tag{3.14}$$

等式右边的函数是具体地算出了的. 由于其形式较复杂,这里不写出公式的细节. 若令 $\Omega_\lambda=0$,(3.14)式显然将会还原到标准模型公式(2.41). 现在需要的是用(3.14)式来拟合实测结果,并相应地定出两个参量 Ω_{m_0}和 Ω_λ. 他们拟合的结果示于图 3.3 中. 这是 Ω_{m_0}—Ω_λ 的平面. 面上任一点代表一种模型,它有确定的预言可与实测比较. 从图上看,在很高的置信度内,只有宇宙常数是大于零的模型才符合他们的测量结果. 这已是一个很重要的结果了. 但是 Ω_{m_0} 和 Ω_λ 的值尚难以确切指出.

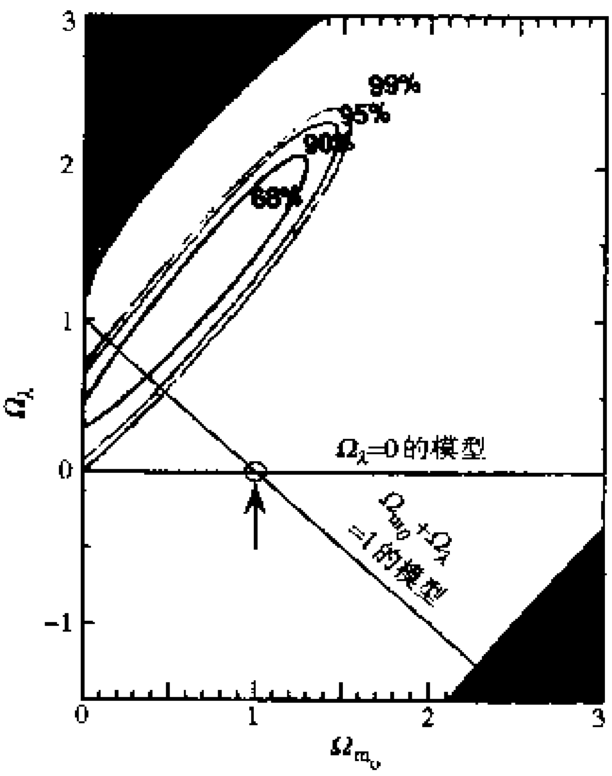


图 3.3 拟合结果画在 Ω_λ—Ω_{m_0} 平面上

从 80 年代以来,有许多原因使理论家相信,包括真空在内的宇宙总密度应当等于 1,即宇宙是平坦的. 我们将在早期宇宙的研究中看到其道理. 在实测上,过去已有较弱的证据表明 Ω_0 在 1 左右. 1999 年,从对微波背景上的温度扰动量的实测分析,才较可靠地证实: Ω_0 确很接近于 1. Perlmutter 等人当时(1998 年)试验地

采用了 $\Omega_0=\Omega_{m_0}+\Omega_\lambda=1$(图 3.3 中的斜线)，定出

$$\Omega_\lambda = 0.65 \pm 0.05. \tag{3.15}$$

相应地有

$$\Omega_{m_0} = 0.35 \mp 0.05. \tag{3.16}$$

由此推知宇宙减速因子约为 $q_0=-0.47$. 他们就这样初步地得出结论：宇宙常数(真空能密度)是正的，今天的宇宙在加速地膨胀. 这称得上是近年来宇宙学观测研究中最重大的成果了.

事情积极的一面很清楚，若这结果是对的，那么标准宇宙模型的结果在定量上是不可靠的，宇宙常数引起的修正是不可忽视的. 但是我们眼下最好保持开放的头脑，而仅把已有的成果看成初步的，有待进一步确证的. 从图 3.2 看，目前所用的数据限于 $z<0.9$ 的星系，距离-红移关系对直线的偏离尚不大，而 Ω_λ 值的推定对这偏离程度又很敏感. 因此，如果最终发现结果有较大的变化，这应不算太意外的事.

四、再论宇宙年龄

前面已分析过，若 $\Omega_\lambda>\Omega_{m_0}$，则宇宙常数将对标准模型的理论结果带来不可忽略的修正. 修正后的结果是否会与事实符合得更好，这当然是我们感兴趣的. 为此这里重新来讨论宇宙年龄问题.

含宇宙常数模型的运动解已经在前面求出(参看(3.7)式). 令 $R(t)/R_0=1$，相应的时刻就是今天的宇宙年龄 t_0. 结果可写成

$$H_0t_0 = G(\Omega_{m_0},\Omega_\lambda), \tag{3.17}$$

其中

$$G(\Omega_{m_0},\Omega_\lambda) = \int_0^1 \frac{\mathrm{d}x}{\sqrt{1+\Omega_{m_0}\left(\frac{1}{x}-1\right)+\Omega_\lambda(x^2-1)}} \tag{3.18}$$

当取 $\Omega_\lambda=0$，这公式就还原为标准模型的年龄公式(2.33).

我们讨论 Ω_λ 对宇宙年龄的影响. 图 3.4 在 Ω_{m_0} - Ω_0($=\Omega_{m_0}+\Omega_\lambda$)平面上画出了函数 G 的等值曲线. G 的值就是 H_0t_0 的值,即用 H_0^{-1} 为单位的宇宙年龄. 从这些等值线看出,对同样的 Ω_{m_0}(即在图上画一条水平线),Ω_λ 越大则预言的宇宙年龄也越大. 因此,若发现实测年龄超过标准模型预言的理论年龄,那也将是宇宙常数大于零的迹象. 前面在标准模型下讨论年龄问题时指出过:要是哈勃常数 H_0 最终被确定在 70 km·s^{-1}·Mpc^{-1}或以上,理论的宇宙年龄将遇到危机. 正的宇宙常数的出现,对缓解这危机将是有利的.

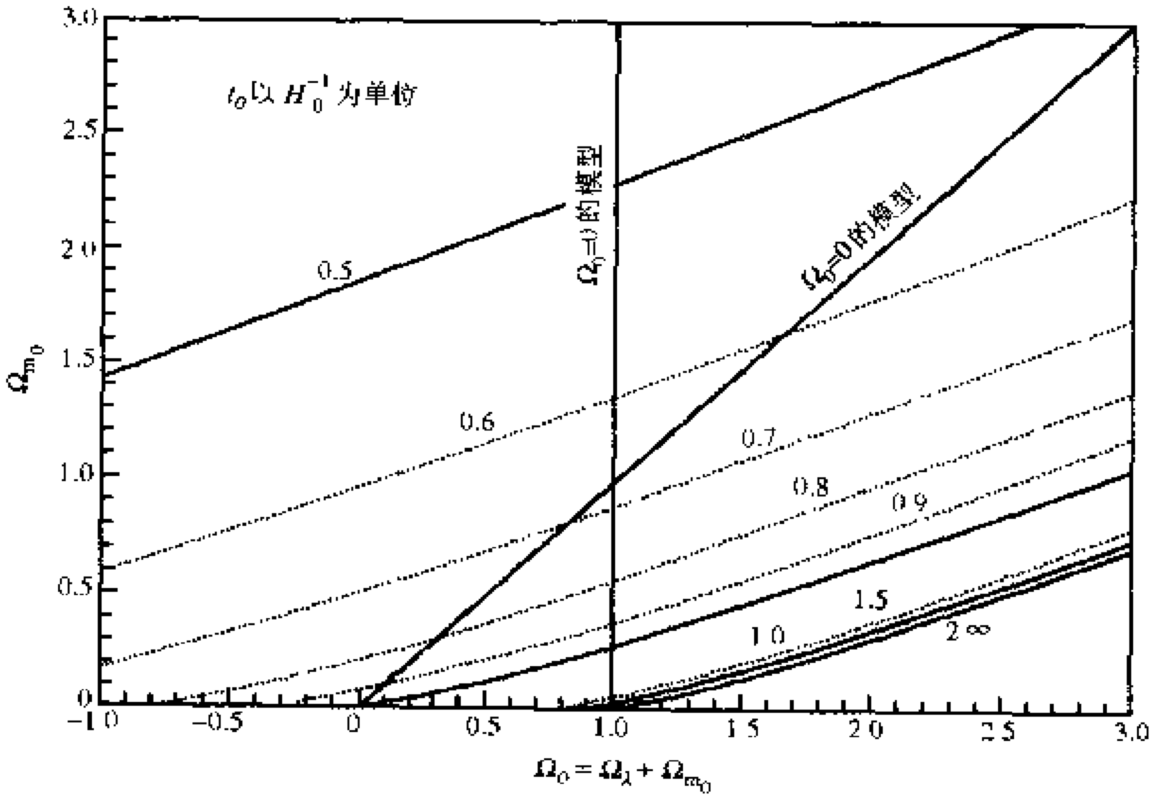


图 3.4　宇宙年龄的等值线

让我们先相信初步定出的真空能密度,即 $\Omega_\lambda=0.65$ 和 $\Omega_{m_0}=0.35$. 由图上读出 G 值为 0.94,即

$$t_0 = 0.94H_0^{-1} = 9.2h^{-1}\,\mathrm{Ga}, \qquad (3.19)$$

其中 h 是以 100 $\mathrm{km \cdot s^{-1} \cdot Mpc^{-1}}$为单位的哈勃常数. 要它与球状星团年龄(12—18 Ga)相符,估出 h 值须在 0.51 到 0.77 之间. 这范围与现今实际测定的哈勃常数范围($h=0.5-0.8$)很重合. 这表明宇宙年龄研究是支持含宇宙常数模型的.

五、宇宙模型的小结

从这两章讨论过的问题已足以使我们意识到,经过近 30 年的广泛研究,虽然还有许多很基本的问题尚未完全解决,但是宇宙学的理论框架已可靠地确立下来了. 在这世纪之交,当人们回顾自然科学的发展时,常把宇宙学理论模型的确立看成最大的成果之一. 但是鉴于前面已谈到的现状,让我们对宇宙学的成就有一个恰如其分的理解.

宇宙学作为物理学的一个分支,它同样是一门实验科学. 理论的正确与否,只能以是否符合事实为准. 可是因为今天的宇宙学测量还没有足够的精度,以可靠地证实或证伪宇宙学的理论预言,所以让我们对这理论作些逻辑分析是有益的.

前面讨论的宇宙理论基于两条基本假设. 一是宇宙学原理. 它说宇宙介质和时空都是均匀和各向同性的. 二是断言广义相对论能作为研究宇宙动力学的出发点. 逻辑地讲,如果前提可靠,那么其推论也一定是正确的. 因此这两条前提是宇宙学理论框架的基础.

近 30 年中观测证实了宇宙学原理为真实宇宙给出了好的近似描写. 这为理论的确立奠定了一块最重要的基石. 所需的另一块基石是广义相对论本身的正确性和它对研究宇宙的适用性. 如果这方面也有了可靠的证据,那么今天的宇宙理论就只会是对的,而不可能是错的. 但是这问题的回答就不那么简单了.

如所周知,广义相对论在建立之初,就有了两个定量上成功的检验,即水星近日点的进动和光线被引力弯曲. 但是它们都是广义

相对论效应较弱的情形. 对特征性的预言如超高密天体和引力波，长期来并没有观测证据. 近二三十年里，对中子星、恒星级黑洞、活动星系核和引力辐射的研究进展，都已为广义相对论的正确性提供了有力的支持. 这实际上也是对宇宙学理论的支持. 无论如何，把广义相对论用于宇宙，它本身存在着 $\lambda=0$ 或 $\lambda\neq0$ 的不确定性. 于是宇宙模型也相应地分成了两支：多数研究采用的 $\lambda=0$ 的标准模型和含宇宙常数的模型.

现在我们已清楚，这两种模型只不过是同一个理论框架下的不同延伸. 前者是后者在 $\lambda=0$ 下的特例. 反过来看，由于观测早已给出宇宙常数不可能太大的限制(参看(3.8))，所以后者只能是前者的修正. 上一章中已说明，标准模型的结果与大量宇宙学事实没有出现明显的矛盾. 至多只有可能出现冲突的初步迹象. 现在又看到，当有冲突出现的初步迹象时，宇宙常数带来的修正将能使冲突得到解决或缓解. 这局面是人们相信宇宙学框架已可靠地确立的重要根据. 但是我们不要把理论框架的已确立理解为标准模型已可靠地被证实. 实际上，宇宙常数是否为零至今尚没有最终的答案，因此理论模型的细节也还没有完全确定. 在理解宇宙学取得的成就时，注意到这一点是十分必要的.

前面已有的讨论还使我们看到，若完全不管中、小尺度上的不均匀，宇宙模型太简单了，所以能够检验的方面并不多. 近几十年里，宇宙学的一个重要进展是把理论模型引申到宇宙演化的早期，即一切星系形成之前的阶段. 这正是在后继的五章中要讨论的内容. 这方向上的成功是宇宙学理论被人们确信的又一重要依据. 有趣的是，宇宙常数是否为零对早期宇宙的研究没有影响. 因为那时实物(或辐射)的密度已很大，宇宙常数的贡献已完全可忽略了. 在这意义上，宇宙早期的研究结果反面少了一点不确定性.

第四章　宇宙的早期

Gamow 首先想到，宇宙早期的面貌并不是今天的缩影. 它曾经没有星系，也没有恒星，而只是由微观粒子构成的高度均匀的普通气体. 时间越早，不仅密度越大且温度也越高. 这就是热大爆炸宇宙理论. 在这早期，宇宙气体的温度曾异常地高，变化幅度也异常地大. 正是这温度的变化使宇宙介质发生了许多次质变，并从而给宇宙演化的研究带来了丰富的物理内涵. 这一章只是作为准备，先对早期宇宙的物理性质作些一般讨论.

一、大爆炸的概念

40 年代末，Gamow 开始引申宇宙膨胀的图景，研究了它演化的早期. 当时的学术界对宇宙膨胀理论本身尚不能接受，因此他的引申被斥之为伪科学. 今天回过头看，当初 Gamow 的思想不仅很深刻，而且其物理基础是相当牢固的.

用宇宙一直在膨胀的观念出发作推理，从物理上讲，若干结论几乎是不可避免地要产生的. 让我们先作些定性分析.

星系只能是宇宙演化的产物

宇宙的膨胀使星系在相互远离. 那么往过去追溯，星系的间距在缩小. 今天星系平均间距仅是星系自身大小的 100 倍不到. 这样就引申出一个推论：当宇宙尺度因子比今天小 2 个量级时，星系

的间距应比星系本身的尺度还小．这是不可能的．它只能意味着那时星系是不会存在的．比方说，若把普通气体压缩到使其中的分子间距比分子本身的尺度还小，那时分子就无法存在．按这样的道理，今天天文学家在研究的一切天体，包括恒星、星系、星系团等，在宇宙早期都是不存在的．它们必是宇宙演化到一定阶段的产物．这推论既不可避免，也很深刻．

星系形成前的宇宙介质是普通气体

那么在一切星系形成之前，宇宙介质以什么状态存在?这就不是一个能由物理规律来推断的问题了．20 世纪初，Jeans 曾提出过自引力不稳定理论．若在一大片均匀气体中有小的密度起伏，气体内的自引力将使小的不均匀放大，直至物质发生局部的结团．按这道理猜想，在各种天体形成之前，宇宙介质应是一大片(无穷大?)带有微小密度起伏的均匀气体．后来由于自引力不稳定性，才把其中小的密度起伏演变成了今天看到的天体．

当然，在理解这想法的合理性时必须注意：它终究是一种猜想，因而是需要由新的观测来证实的．这里提前指出，后来微波背景辐射的研究直接地证实了它．这样人们清楚了：早期宇宙非常简单，它是一片没有宏观结构的均匀气体．这实际上是宇宙学原理在早期的体现．

时间越早则宇宙越热

在均匀宇宙中，各部分的温度也是相同的．在宇宙膨胀时，各部分气体间不会有热量的流动，即它作的是绝热膨胀．气体的绝热膨胀要降温，因此早期宇宙不仅密度越早越大，其温度也越高．这又是关于早期宇宙的一个必然的推论．

若一直往前追溯下去，必在某有限时刻，宇宙温度将随密度一起趋向无穷大．前面已指出过，宇宙的膨胀过程有一个起始时刻是不可避免的结果，也是许多人难以接受的结果．人们把宇宙有"创生"看作是科学与宗教的妥协，因此拒绝接受 Gamow 对创生后不久的宇宙的研究．历史上，正是反对 Gamow 理论的人讥讽地把它

称之为宇宙的大爆炸理论. 无论如何,今天的宇宙学研究已冲破贫乏哲学的困扰,而深深地进入了大爆炸后的第一秒之内. 大家也已普遍地意识到宇宙确实存在着怎么创生的问题. 但是它依然是一个物理问题而不是宗教或哲学问题.

顺便讲,大爆炸宇宙学这名称虽然今天已被正面地沿用了下来,但它却有误导性,好像大爆炸理论事先假定了宇宙是有限的. 我们已讨论过,宇宙学并没有把这当前提,而且宇宙的有限或无限也不是可以随便假设的. 过分地从字面上来理解大爆炸的含义是一种误解. Gamow 理论只是指出,宇宙早期演化是从密度和温度都非常高的状态开始的.

原子和分子是宇宙演化的产物

我们知道,分子的解离能远小于 1 eV,氢原子的解离能则等于 13.6 eV. 因此,当气体的温度 T 超过 105 K,即其组分粒子的平均热动能① kT 超过 10 eV,热碰撞将导致分子和原子的解离. 那时整个宇宙介质将处于等离子体状态. 主要组分粒子是原子核、自由电子和光子等. 这样看来,原子和分子也是宇宙演化的产物,而不是从来就存在的.

Gamow 首先发现,在宇宙降温至中性原子形成时,宇宙气体中将形成背景光子. 这些背景光子应遗留至今天,从而是可以用观测来寻找的. 这是下一章中要讨论的专题.

连化学元素也是宇宙演化的产物

我们知道,原子核也是复合粒子. 它们是由核子(质子和中子的统称)组成的. 核物理告诉我们,每一个核子在原子核中的结合能是 1 MeV 的量级. 这样,给原子核以 1 MeV 的能量,就能够把其中一个核子解离出来. 那么当宇宙温度超过 1 MeV(即 10^{10} K)时,不断的热碰撞将把原子核完全瓦解. 在这样高温的宇宙中,原

---

① 在自然单位制中令 Boltzmann 常数 $k=1$,因此粒子平均热动能 $E=T$. 这样就能用 eV 来量度温度,$T=1$ eV 相当于 104 K. 宇宙学中用 eV 来表示温度的好处是直观地告诉了我们气体中组分粒子的热动能.

子核是不能存在的.宇宙气体将由质子、中子、电子、光子和其他的基元粒子组成(我们将把它叫做粒子气体).若把等离子态叫物质的第四态,那么这应是物质的第五态.

化学元素的性质是由其原子核决定的.宇宙的早期曾没有原子核,当然也就没有元素.连化学元素那样基本的物质要素都是在宇宙演化中产生的,这认识无疑很深刻.

上面的定性分析使我们看到了早期宇宙的两面.首先它非常简单,只不过是温度在逐渐降低的均匀气体.同时又看到,在温度降低的过程中,它蕴涵着非常丰富的物理变化.结合这两面,今天显得很复杂的宇宙怎么从很简单的状态一步一步地演化出来的图景已露端倪了.

二、辐射为主的早期

我们把宇宙物质的组分粗分为实物和辐射两大类.今天的宇宙以实物为主是一个事实.早期宇宙以辐射为主,却是物理推论.

常规地想,今天宇宙中的实物就是由分子和原子组成的普通物体.事实上很可能其主要部分不是这样的普通物质.这是后面要讨论的非重子暗物质问题.现在只先讨论实物的总密度.虽然很大一部分实物不发光,但是实物总密度的量级是已清楚的.取用 $\Omega_m=0.3$, $h=0.65$ 来分析,它是

$$\rho_m \approx 2\times10^{-27}\,\mathrm{kg/m^3} = 1\times10^{9}\,\mathrm{eV/m^3}. \qquad (4.1)$$

它相当于每立方米内有一个氢原子的质量.若把实物全当作核子,其数密度是

$$n_N \approx 1/\mathrm{m^3}. \qquad (4.2)$$

辐射的主要来源是背景光子(先不考虑中微子).它的粒子数密度是(参看下一章)

$$n_\gamma \approx 400/\mathrm{cm^3}. \qquad (4.3)$$

因此论粒子数，实物与辐射之比[1] η 是

$$\eta = n_N/n_\gamma \approx 3 \times 10^{-9}, \tag{4.4}$$

即辐射粒子约多 9 个量级. 但是辐射光子的平均能量 E_γ 很小，它仅是

$$E_\gamma \approx 3 \times 10^{-4}\,\text{eV}. \tag{4.5}$$

与核子的静能（~1 GeV）相比，它要小 13 个量级. 密度是粒子数密度与粒子质量（即能量）的乘积. 因此，辐射密度反而比实物密度小 4 个量级，即

$$\rho_\gamma/\rho_m \approx 10^{-4}. \tag{4.6}$$

这就是今天宇宙以实物为主的具体含义.

然后我们分析早期宇宙必以辐射为主的原因. 看宇宙尺度因子 $R=10^{-4}R_0$ 的早期. 在宇宙膨胀中，实物粒子和光子的数密度都要减小，且都与 R^3 成反比，因此两者之比 η 是不变的. 在那时，光子数仍比核子数多 9 个量级. 核子的质量仍来自静质量，它没有变化，但是光子的平均质量变大了. 光子的动质量来自热运动，因而它与温度成正比. 第四节中将证明，早期宇宙的温度与尺度因子成反比，即有

$$E_\gamma \approx T_\gamma \propto R^{-1}. \tag{4.7}$$

这样看来，在 R 降了 4 个量级的早期，每一光子的质量（即能量 E_γ）大了 4 个量级，即达到 3 eV. 那时光子已只比核子轻 9 个量级，而粒子数多 9 个量级. 于是辐射与实物的密度几乎相同. 这叫辐射与实物的等量时刻. 再往前追溯，ρ_γ 就超过了 ρ_m，于是形成了辐射为主的早期.

粗略讲，以辐射为主的早期是 $R<10^{-4}R_0$ 的阶段. 折合到宇宙时间，它大体是宇宙演化的最初 1 万年. 对于已演化了约 150 亿年的宇宙讲，这只是一个很短的初始阶段. 以前在讨论宇宙年龄时

---

① 这把实物全当成了核子. 核子实际上可能只占 10%，因此真正的 η 值要小一个量级.

$$\rho = \frac{\pi^2}{30} g^* T^4, \tag{4.10}$$

其中

$$g^* = g_B + \frac{7}{8} g_F. \tag{4.11}$$

这里用 g_B 与 g_F 分别代表各玻色或费米组分的自旋自由度数之和. g^* 常被称混合气体的总(自旋)自由度数,但它不简单地是 g_B 与 g_F 之和,因为玻色子与费米子的贡献有一点小的差别.

从(4.10)式看,密度正比于温度的四次方,这是辐射气体的主要特征. 在温度确定后,我们还需要知道自然界有多少种 $m<T$ 的基元粒子,以便把总自由度数 g^* 算出来. 这自然需要一些粒子物理的知识(参看附录Ⅱ). 但是要弄清这一点并不难. 作为后面有用的例子,看 $T=1$ MeV 的情形. 从附录Ⅱ中的粒子表可知,$m<1$ MeV 的粒子只有光子、三代中微子,电子以及它们的反粒子(光子的反粒子就是自身). 光子是自旋为 1 的玻色子,其自旋自由度是 2. 正反中微子都是自旋为 1/2 的费米子,每种只有 1 个自旋自由度. 正反电子也是自旋为 1/2 的费米子,其自旋自由度各为 2. 这样就可算出 $g_B=2$,$g_F=10$,因此有 $g^*=10.75$. 对一般情况,我们只需有一个粗略的印象:当 $T<1$ GeV,g^* 值接近为 10;当 $T\gg 1$ GeV,强作用粒子将以夸克状态存在,按粒子物理的标准模型估算,相应的 g^* 值约为 100. 由此看到,g^* 值虽然与温度有关,但随 T 的变化十分缓慢. 因此在后面的讨论中,它将被当作常数看待.

统计物理给出,辐射组分的压强与密度有很简单的比例关系:

$$P = \rho/3, \quad \text{不分玻色子或费米子}. \tag{4.12}$$

由于任何组分的压强与密度关系都与光子气体一样,因此辐射气体的总压强与总密度也满足(4.12). 在第二章中曾用这样的关系证明了(参看(2.15))

$$\rho R^4 = \text{常数}. \tag{4.13}$$

现在看来,整个早期这关系将都成立. 把它与密度 ρ 正比于 T^4 结

$$\rho = \frac{\pi^2}{30} g^* T^4, \tag{4.10}$$

其中

$$g^* = g_B + \frac{7}{8} g_F. \tag{4.11}$$

这里用 g_B 与 g_F 分别代表各玻色或费米组分的自旋自由度数之和. g^* 常被称混合气体的总(自旋)自由度数,但它不简单地是 g_B 与 g_F 之和,因为玻色子与费米子的贡献有一点小的差别.

从(4.10)式看,密度正比于温度的四次方,这是辐射气体的主要特征. 在温度确定后,我们还需要知道自然界有多少种 $m<T$ 的基元粒子,以便把总自由度数 g^* 算出来. 这自然需要一些粒子物理的知识(参看附录Ⅱ). 但是要弄清这一点并不难. 作为后面有用的例子,看 $T=1$ MeV 的情形. 从附录Ⅱ中的粒子表可知,$m<1$ MeV 的粒子只有光子、三代中微子,电子以及它们的反粒子(光子的反粒子就是自身). 光子是自旋为 1 的玻色子,其自旋自由度是 2. 正反中微子都是自旋为 1/2 的费米子,每种只有 1 个自旋自由度. 正反电子也是自旋为 1/2 的费米子,其自旋自由度各为 2. 这样就可算出 $g_B=2$,$g_F=10$,因此有 $g^*=10.75$. 对一般情况,我们只需有一个粗略的印象:当 $T<1$ GeV,g^* 值接近为 10;当 $T\gg1$ GeV,强作用粒子将以夸克状态存在,按粒子物理的标准模型估算,相应的 g^* 值约为 100. 由此看到,g^* 值虽然与温度有关,但随 T 的变化十分缓慢. 因此在后面的讨论中,它将被当作常数看待.

统计物理给出,辐射组分的压强与密度有很简单的比例关系:

$$P = \rho/3, \quad \text{不分玻色子或费米子.} \tag{4.12}$$

由于任何组分的压强与密度关系都与光子气体一样,因此辐射气体的总压强与总密度也满足(4.12). 在第二章中曾用这样的关系证明了(参看(2.15))

$$\rho R^4 = \text{常数}. \tag{4.13}$$

现在看来,整个早期这关系将都成立. 把它与密度 ρ 正比于 T^4 结

$$\rho = \frac{\pi^2}{30} g^* T^4, \tag{4.10}$$

其中

$$g^* = g_B + \frac{7}{8} g_F. \tag{4.11}$$

这里用 g_B 与 g_F 分别代表各玻色或费米组分的自旋自由度数之和. g^* 常被称混合气体的总(自旋)自由度数,但它不简单地是 g_B 与 g_F 之和,因为玻色子与费米子的贡献有一点小的差别.

从(4.10)式看,密度正比于温度的四次方,这是辐射气体的主要特征. 在温度确定后,我们还需要知道自然界有多少种 $m<T$ 的基元粒子,以便把总自由度数 g^* 算出来. 这自然需要一些粒子物理的知识(参看附录Ⅱ). 但是要弄清这一点并不难. 作为后面有用的例子,看 $T=1$ MeV 的情形. 从附录Ⅱ中的粒子表可知,$m<1$ MeV 的粒子只有光子、三代中微子,电子以及它们的反粒子(光子的反粒子就是自身). 光子是自旋为 1 的玻色子,其自旋自由度是 2. 正反中微子都是自旋为 1/2 的费米子,每种只有 1 个自旋自由度. 正反电子也是自旋为 1/2 的费米子,其自旋自由度各为 2. 这样就可算出 $g_B=2$,$g_F=10$,因此有 $g^*=10.75$. 对一般情况,我们只需有一个粗略的印象:当 $T<1$ GeV,g^* 值接近为 10;当 $T\gg1$ GeV,强作用粒子将以夸克状态存在,按粒子物理的标准模型估算,相应的 g^* 值约为 100. 由此看到,g^* 值虽然与温度有关,但随 T 的变化十分缓慢. 因此在后面的讨论中,它将被当作常数看待.

统计物理给出,辐射组分的压强与密度有很简单的比例关系:

$$P = \rho/3, \quad \text{不分玻色子或费米子.} \tag{4.12}$$

由于任何组分的压强与密度关系都与光子气体一样,因此辐射气体的总压强与总密度也满足(4.12). 在第二章中曾用这样的关系证明了(参看(2.15))

$$\rho R^4 = \text{常数}. \tag{4.13}$$

现在看来,整个早期这关系将都成立. 把它与密度 ρ 正比于 T^4 结

合起来，得到

$$RT = \text{常数}. \tag{4.14}$$

这就是上节中用到过的结果：R 与 T 成反比，R 每增大一个量级，则温度降低一个量级.

组分粒子的数密度也常是宇宙学问题所关心的. 相应的结果是

$$n = \begin{cases} \frac{\zeta(3)}{\pi^2} g T^3, & \text{对玻色子}, \\ \frac{3}{4} \cdot \frac{\zeta(3)}{\pi^2} g T^3, & \text{对费米子}, \end{cases} \tag{4.15}$$

其中 $\zeta(3)=1.202\cdots$是宗量为 3 的 Riemann ζ 函数值，g 仍是自旋自由度数. 这公式告诉我们，粒子数密度 n 与 T^3 成正比. 结合(4.14)知，n 与 R^3 成反比. 这意思是说，如果没有非热过程发生，那么粒子数密度与共动体积成反比. 宇宙膨胀使某一群粒子所占的体积增大，其数密度则反比地降低. 这是一个很自然的结果.

四、早期膨胀的动力学

不管早期或晚期，宇宙的膨胀总应满足 Friedmann 方程. 先把实物与辐射都计入，连宇宙常数也考虑在内，它可写成(参看(3.1))

$$\left(\frac{\dot{R}}{R}\right)^2 = -\frac{k}{R^2} + \frac{8\pi G}{3}(\rho_m + \rho_\gamma + \rho_{eff}). \tag{4.16}$$

注意右边各项与 R 的比例关系很不一样：

真空项	$\rho_{eff} \propto R^0$,	(4.17)
曲率项	$k/R^2 \propto R^{-2}$,	(4.18)
实物项	$\rho_m \propto R^{-3}$,	(4.19)
辐射项	$\rho_\gamma \propto R^{-4}$.	(4.20)

往早期追溯，辐射项增长得最快. 这正是早期以辐射为主的原因.

今天宇宙中的实物项、真空项和曲率项都同量级，这是事实. 在 R 减小时，后两项比实物项增长得慢，因此追溯到实物密度已可忽略的早期，曲率项和真空项更可忽略了. 这样在早期宇宙动力学方程中，右边只要保留辐射项就够了. 于是它被简化为

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G}{3}\rho, \tag{4.21}$$

其中的 ρ 仅指辐射组分的密度，但为方便而不写下标了.

对宇宙早期，尺度因子 R 的含义已很不直观. 考虑到 R 与温度 T 成反比，我们宁肯把动力学方程化成温度 T 的方程. 由 $RT=$ 常数知

$$\dot{R}/R = -\dot{T}/T. \tag{4.22}$$

把密度与温度的关系(4.10)代入，动力学方程(4.21)化成了

$$\dot{T}^2 = \frac{4\pi^3 G}{45}g^* T^6. \tag{4.23}$$

近似地把 g^* 当常量，易于把这方程解出，得到

$$T = At^{-1/2}, \tag{4.24}$$

其中

$$A = \left(\frac{45}{16\pi^3 G g^*}\right)^{1/4}. \tag{4.25}$$

系数 A 对 g^* 的大小不敏感，因此把它当常数是合理的近似.

这样解出的是自然单位制下的结果，实用上常愿意用秒做时间单位. 折合后算出，宇宙年龄为 1 s 时的介质温度约为 1 MeV. 记住这结果，其他时刻的温度就可按比例来推算. 宇宙年龄增大 2 个量级，则温度降低 1 个量级. 于是早期宇宙中任一时刻的温度就大致清楚了.

最后说明一下密度随时间的变化. 粗略的结果是这样：宇宙年龄为 1 s 时的密度是 10^6 g/cm^3，其他时刻可用 ρ 与 t^2 成反比来推算.

五、宇宙演化简史

在知道宇宙温度及密度随时间的变化后，宇宙演化的一个简史就能编出来了.

现今的宇宙学以广义相对论为基础，说明引力场须是经典场.因此对过高的温度，这模型是不适用的.那时引力作为量子场的效应将很显著.必须考虑引力量子性的特征能量是 $G^{-1/2}$，被叫做普朗克能量.由此估出经典宇宙的“物理”起点是

$$T \sim G^{-1/2} = 10^{19}\,\text{GeV}. \tag{4.26}$$

此前的宇宙须用量子引力理论研究.与这能量相对应的时间为

$$t \sim G^{1/2} = 10^{-43}\,\text{s}, \tag{4.27}$$

叫普朗克时间.这应该被当作经典宇宙的开端.

从 10^{-43} s 到 10^{-4} s，宇宙温度从 10^{19} GeV 降至 0.1 GeV. 按粒子物理的标准模型分析，这最早阶段的宇宙气体应由夸克、轻子和规范粒子等所组成.因为粒子物理对能量远高于 10^3 GeV 的规律尚不确切掌握，所以宇宙的这段历史像远古史，还难以说得很清楚.按现有研究，这阶段应发生过的过程有真空相变引起的暴胀，正反重子不等量的产生和冷暗物质的形成等.它们都对宇宙学有重大价值，但这些过程的研究尚带有试探性，没有很肯定的结果.人们把对这阶段的研究称为粒子宇宙学.

约在 $t=10^{-4}$ s，宇宙温度降到了 0.1 GeV，密度降至 $10^{15}\,\text{kg/m}^3$. 从温度和密度条件看，这时刻前后应会发生从夸克到强子的相变.此后的宇宙气体中才有了质子和中子.由于正反夸克对已大量地湮灭，所以剩下的只有核子而没有反核子.其数密度已比光子低了 9 至 10 个量级.再往后就该发生原子核的合成了.

至 $t>1$ s，温度降至 1 MeV 以下.宇宙进入了核物理的能量范围.人们对相应的物理规律已相当清楚，因此宇宙学问题的研究开始有了可靠的基础.原初的核合成是在 $t=3$ 分至 1 小时内发生

的.理论上对所发生的过程已有了很详尽的研究.这过程结束,宇宙中开始有了化学元素,主要是氢(包括氘)和氦,还有极少的锂、铍、硼.不过它们还是处于等离子状态.下面该等待的是中性原子的形成.

从 $t \approx 10^{12}$ s $\approx 10^{4}$ a 起,宇宙温度降到了 10 eV 以下.这是原子物理的能量范围,其物理规律更是十分清楚.约在温度降至 0.3 eV 时,原子核和自由电子开始结合成了中性原子.于是宇宙介质成了很普通的中性原子气体.原来存在着的热光子从此失去了热碰撞对象,作为背景光子存留了下来.

观测研究已发现,那时的均匀气体上有小的密度起伏.这起伏将按自引力不稳性的规律演化,最后将形成结团状的天体.这过程发展得很慢.最早的星系约形成于宇宙年龄为 1×10^{9} 年前后.恒星则是星系物质进一步碎裂的产物.宇宙就这样逐渐演变成了今天看到的模样.

这是对宇宙演化的一个极简单的陈述,或者说是宇宙演化史的梗概.对其中各个环节,人们了解的清楚程度很不相同.大致地说,宇宙学就像考古学一样,越早越难以很清楚.因此本书后面的专题讨论将逆着时间方向进行,先讨论背景辐射问题,再逐步往前.星系形成虽发生得很晚,但它的研究须以各方面的成果做基础,只能放在最后讨论.

六、粒子的退耦

我们一直把膨胀着的宇宙气体看成是热平衡的.为实现热平衡,必须要有足够频繁的热碰撞.当宇宙年龄很短而膨胀又很快,气体能否有足够频繁的碰撞以保持热平衡?这是一个需要研究的问题.我们将在这里讨论.讨论的背后有一个更重要的动机,那就是由失去热平衡的条件而引申出粒子退耦的概念.这是研究早期宇宙所不可缺少的准备概念.

为了简单，考虑单组分气体. 令 Γ 为每一粒子在单位时间内的碰撞次数，则有

$$\Gamma = n\langle v\sigma(v)\rangle, \tag{4.28}$$

其中 n 是靶粒子的数密度，v 和 $\sigma(v)$ 分别是相对速度和碰撞截面，$\langle\cdots\rangle$ 代表对速度分布作平均. 热平衡下的平均值与气体温度有关，说明碰撞频率 Γ 取决于气体的温度、密度和相互作用的性质.

对膨胀着的宇宙，问题在于：尺度因子 R 有显著变化的时间间隔 Δt 内能否有足够的碰撞次数以保持热平衡. 用 $\Delta R \sim R$ 作显著变化的标志，考虑到膨胀速率 $H = \Delta R/\Delta t/R$，则有

$$\Delta t \sim 1/H, \tag{4.29}$$

其中 H 是任意时刻的膨胀速率. 把每一粒子有频繁碰撞的条件写成 $\Gamma\Delta t \gg 1$. 结合(4.28)，这等于要求

$$\Gamma \gg H. \tag{4.30}$$

那就是说，若粒子的碰撞率远超过宇宙膨胀率，则气体中能有足够的碰撞来维持热平衡. 这是基本结果. 由此引申出热平衡有被破坏的可能.

随着宇宙的膨胀，Γ 和 H 都是在降低的. 因此两者之间有一种竞争关系. 若碰撞率降得快，则条件(4.30)总将被破坏. 这种粒子就会退出热平衡，并进一步完全失去热碰撞的机会. 后一步就是粒子的退耦. 让我们以仅参与弱作用的中微子为例来分析.

按弱作用的性质算出粒子碰撞率为

$$\Gamma \sim G_F^2 T^5, \tag{4.31}$$

其中 G_F 是弱作用的费米耦合常数. 我们已经知道，早期宇宙的膨胀率为

$$H \sim t^{-1} \sim T^2. \tag{4.32}$$

它是一个粒子碰撞率下降更快的现实例子. 当用 MeV 为能量单位，费米耦合常数 G_F 的大小接近是 1. 由此看出，当宇宙温度 $T \gg 1$ MeV 时，热碰撞足够频繁的条件是满足的. 在 $T = 1$ MeV 后将失去热平衡. 然后让我们一般地论证：在 $\Gamma \approx H$ 之后，这种粒子已

实际上完全失去了碰撞机会.

设某种粒子在温度为 T_d(时间为 t_d)时达到 $\Gamma \approx H$. 从这时刻到无穷远的将来,粒子的总碰撞次数 N 为,

$$N = \int_{t_d}^{\infty} \Gamma(t')\mathrm{d}t'. \tag{4.33}$$

利用 $H = -\dot{T}/T$,并改用温度 T 代替时间 t(记得 T^2 反比于 t),上式化作

$$N = \int_0^{T_d} \frac{\Gamma(T')}{H(T')} \frac{\mathrm{d}T'}{T'}. \tag{4.34}$$

把粒子碰撞率写成 $\Gamma \propto T^n$,并令 $n>2$ 以代表碰撞率比膨胀率下降得更快的情形. 上述积分易于积出,得到

$$N = \frac{1}{n-2} \frac{\Gamma(T_d)}{H(T_d)} \sim 1. \tag{4.35}$$

这结果说明,粒子从那时起到无穷远的将来,还将有约一次的碰撞机会.

我们从这讨论中得到的感兴趣的结论是:对于碰撞率下降较快的粒子,宇宙温度降至 T_d 后,它实际上已在混合气体中成为无碰撞组分了. 这就是该粒子在宇宙中退耦的概念. 意思是从此它与其他组分失去了(引力之外的)物理联系. 退耦后的粒子将孤立地存在下去,直至今天和很遥远的将来. 它因此被称为背景粒子. 通常讲的背景辐射,指的就是退耦后的光子组分. 同样,退耦后的中微子就成了背景中微子. 它在今天也应该仍存在着,只是不易被发现罢了. 所有退耦后的粒子都是早期宇宙留下来的珍贵遗迹. 宇宙学研究对它们特别感兴趣,原因正在于此.

七、宇宙的视界

在第一章里谈基本事实时已指出过,宇宙中可被我们观测研

究的部分总是有限的．由于光速有限，在年龄又有限[①]的(经典)宇宙中，不管客观上它本身为有限或无限，观测者只有一个有限大的可观测范围，这完全是不可避免的推论．这范围的边界被称为宇宙对该观测者的视界．让我们讨论视界的概念，并估算视界的大小．

图 4.1 给出了视界概念的图解．其中横轴表示随动的径向坐标，纵轴代表时间．设观测者处于坐标原点，而光源的坐标为 r．宇宙的膨胀使两者距离变远是通过 $R(t)$ 增大体现的，它们的坐标保持不变．图上的实线画出 $r=r_h$ 处的光源在 $t=0$ 时发出的光，它正好在 t 时刻到达原点．容易看出，更近的光源总有一个时刻 t_1，这时发出的光也在 t 时到达原点．但是对更远的光源，它在任何时刻发出的光都来不及在 t 时到达．这意味着处于原点的观测者原则上看不到更远的物体．因此，这 r_h 反映了 t 时刻的可观测范围，即视界的大小．

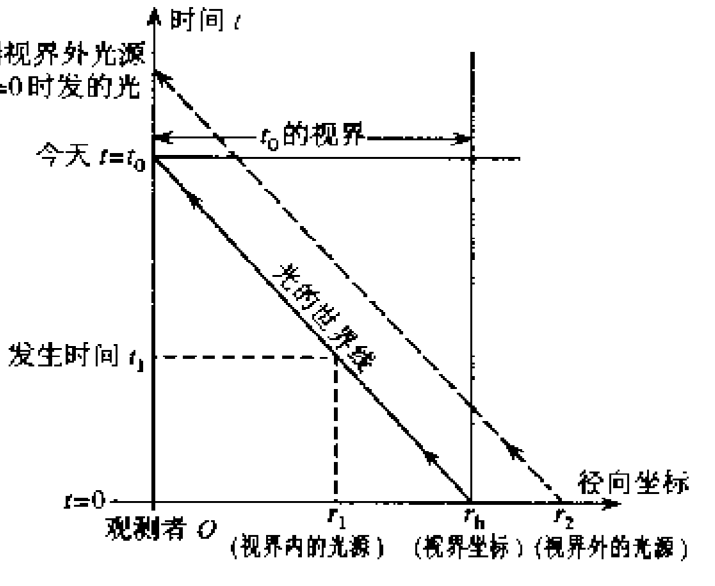


图 4.1　今天的视界

① 指允许光在宇宙中传播的时间是有限的．

r_h 与 t 的关系可具体算出. 广义相对论要求光的传播满足 $ds=0$. 在 Robertson-Walker 度规中,这要求体现为

$$\frac{dt}{R(t)}=\pm\frac{dr}{\sqrt{1-kr^2}}. \qquad (4.36)$$

光向外传播取+号,向中心传播取-号. 这样作积分后,就得到了视界坐标 r_h 与时间 t 的关系,

$$\int_0^t\frac{dt}{R(t)}=\int_0^{r_h}\frac{dr'}{\sqrt{1-kr'^2}}. \qquad (4.37)$$

注意随动径向坐标 r 不直接代表距离. t 时坐标 r 处的点与原点的距离定义为

$$l(t)=R(t)\int_0^r\frac{dr'}{\sqrt{1-kr'^2}}. \qquad (4.38)$$

人们常把它叫固有距离. 结合(4.37)和(4.38),就能把视界的固有大小算出. 它是

$$l_h(t)=R(t)\int_0^t\frac{dt'}{R(t')} \qquad (4.39)$$

看来为了得到视界的大小,只须先把 $R(t)$ 求出. 这是前面已做到了的事.

对早期宇宙,用 $R=At^{-1/2}$ 代入,立即有

$$l_h(t)=2t. \qquad (4.40)$$

对后期宇宙,$R(t)$的形式与曲率因子 k 有关. 对于讨论视界,要的仅是量级,因此用 $k=0$ 的情形做代表即可. 把 $R\propto t^{-2/3}$代入,得到

$$l_h(t)=3t. \qquad (4.41)$$

于是任何时刻视界的大小都大体知道了. 这些结果都是在自然单位制下写出的. 若改用普通单位制,只要把 t 换成 ct 就行. 视界的大小与时间成正比,这是不意外的结果. 多少有些意外的是公式中出现了比例因子 2 或 3. 它是宇宙空间在膨胀的后果.

在估算今天的视界大小时,又可以把辐射为主的早期忽略. 再折中地取宇宙年龄 $t_0=150$ 亿年,用(4.41)式估出

$$l_h(t_0) \approx 1 \times 10^4 \text{ Mpc}. \tag{4.42}$$

如已说过，它仅比超星系团或大空洞（的线度）大几百倍. 这或许比我们事先想像的要小. 这里有一点须注意，天文上测到的光度距离比这里定义的固有距离放大了 $1+z$ 倍，z 是观测对象的红移量. 因此，当发现某星系的光度距离超过视界值，并不表示它处在可观测范围之外，而只表明它的红移很大.

视界的概念除描述可观测范围外，它还反映了宇宙中可能有因果联系的区域大小. 若两个地方的固有距离超过视界值，它们之间不可能有任何信息交流，因此两处发生的事情就不可能有因果关系. 这是一个基本概念. 把它用于宇宙，却出现了令人困惑的疑问：宇宙学原理假定在同一时刻，一切地方的物质密度都一样. 若两地点的固有距离超过视界，物理上就没有任何机制能造成它们密度的相等. 这问题被人们称为视界疑难. 我们将在后面讨论它.

第五章　微波背景辐射

宇宙背景辐射的存在是大爆炸理论的主要预言.它在 1965 年被发现,这是宇宙学成功史上最珍贵的一页.后来,关于背景辐射性质的一切理论预言都陆续得到了观测的证实.这使学术界普遍地接受了大爆炸理论.至今对背景辐射的研究仍在进展中,但动机已不再是检验理论,而是想从中获取更多关于早期宇宙的信息.

一、背景光子的形成

我们从宇宙年龄为 10^5 年时讨论起.当时介质温度约为 1 eV.以氢和氦为主的气体是电离的.这电离气体的基本组分是氢核、氦核、自由电子和光子.为便于讨论,下面把它简化为纯氢的等离子气体.

自然,氢原子核(即质子)和自由电子能结合为中性原子,

$$p + e^- \to H + \gamma, \tag{5.1}$$

但是这过程是可逆的.当能量超过 13.6 eV 的光子与氢原于相碰,将使它还原成质子和自由电子.记得光子数目比氢核多 9 至 10 个量级.在 $T=1$ eV 时,足以使原子电离的光子尚太多,结果是正过程与逆过程达到统计平衡.从理论上估算,这情况下中性氢与电离氢之比微不足道.所以人们认为这时宇宙中还没有中性原子.

在气体温度随宇宙膨胀而再降低,能量超过 13.6 eV 的光子数将指数地减少,上述逆过程很快就实际地失效了.这才发生气体

从电离状态向中性原子状态的转化.

理论上要定量地处理这转化过程没有困难.问题在于计算结果依赖于两个我们并不准确知道其值的参量.一是气体中核子数与光子数之比 n_N/n_γ,我们仅知道它的值在 10^{-10}到 10^{-9}之间.另一是今天宇宙物质的总密度 Ω_0. 它的值在 0.2 到 1.0 之间.这两个因素给结果带来了一些不准确性.

粗略地讲,当宇宙年龄为 4×10^5 年时,温度降到 $T=0.3$ eV,那时中性氢原子已在气体中占到 90%. 人们常把它叫原子的复合时刻.此后电离氢与中性氢的比例仍在降低.直到电离氢只占 10^{-5},自由电子失去了与氢核相碰的机会,转化过程才趋向停止,这就是宇宙中最初的中性原子的产生过程.然后我们把注意力转向气体中的光子组分的命运.

当气体处于电离状态时,光子是一个热平衡的组分.光子的热碰撞主要是通过它与自由电子间的 Thomson 散射来实现的. Thomson 散射截面是一个与速度无关的常数,其值为 $\sigma_{Th}=6.65\times10^{-25}$ cm^2. 光子速度 $c=1$ 也是常数.因此按(4.28),光子的热碰撞率为

$$\Gamma_\gamma = n_e c\sigma_{Th} \propto n_e. \tag{5.2}$$

这里的 n_e 是自由电子的数密度.在原子中性化前,因有大量电子处于自由状态,所以无疑有 $\Gamma_\gamma\gg H$. 那时光子有足够频繁的碰撞以维持热平衡.在原子中性化的后期,绝大部分电子束缚在中性原子内,自由电子的数密度的迅速降低,使光子的碰撞率急剧减少,最后它完全失去了碰撞机会.这就是光子在宇宙介质中的退耦.退耦后的光子与气体中的其他组分粒子(中性氢)已几乎没有相互作用,因而被称为背景光子,或背景辐射场.所以说,中性原子的形成和光子的退耦是同一过程中的两个方面.

在光子退耦后,它当然将在宇宙中一直存在下去,直到今天,所以这一个理论结果是可以用实测来检验的.若天文观测否定它的存在,那将是对大爆炸理论的致命打击.反之,若发现了背景光

子，那就是对理论的决定性的支持．因此这种实测研究是对大爆炸理论的严峻考验．

二、观测性质

观测到什么样的信号才能说明发现了背景光子？这是要在检验前必须弄清的基本问题．

为了对宇宙学尺度上的观测有直观的认识，下面完全不成比例地画出了一个整个观测宇宙的示意图（图 5.1）．观测者处在中心位置上．

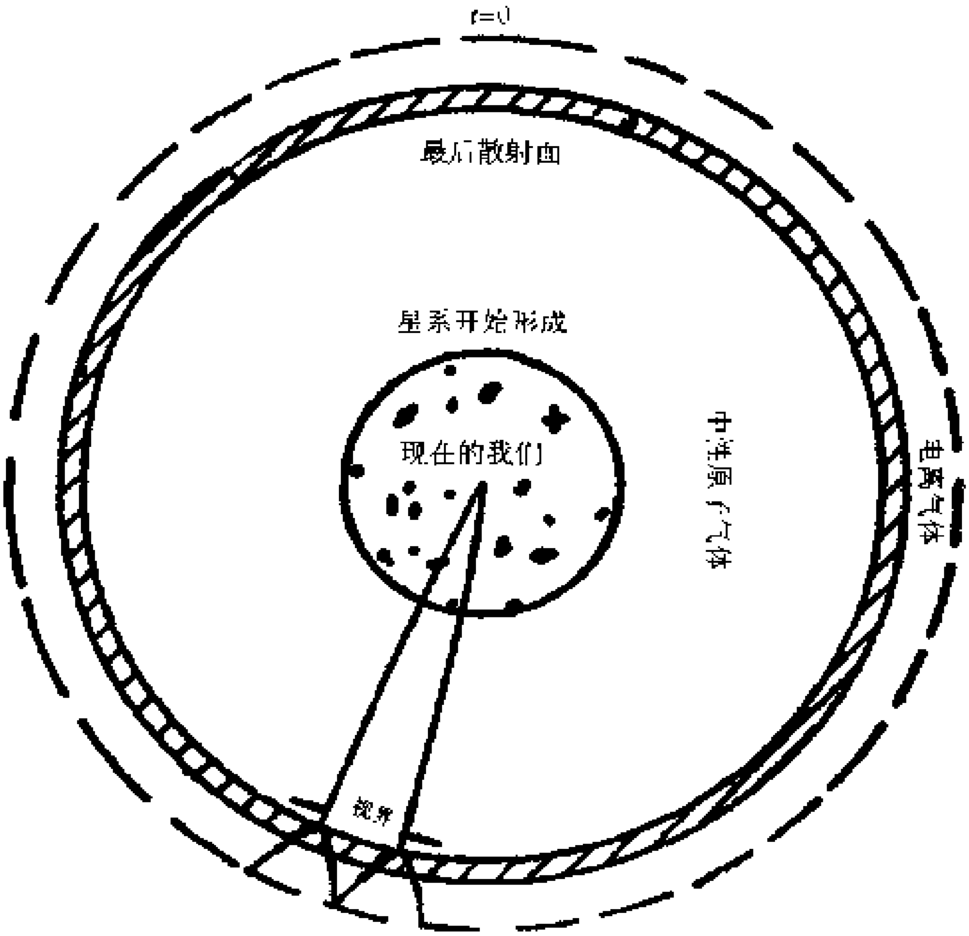


图 5.1　观测宇宙图

我们已经理解，径向坐标为 r_1 处的天体在 t_1 时所发的光才正好被我们今天(t_0 时)观测到，而且 r_1 越大则 t_1 越小. 这是理解观测宇宙图的基础. 这是光速有限性造成的. 按这道理，观测宇宙图不是等时图. 在越远的地方，我们是在观测那里越早的历史情景. 反之，那里较晚(或今天)的情景是观测不到的，因为所发的光还没有到达. 这也就是说，通过对宇宙空间纵深的观测，看到的是宇宙演化的历史.

观测宇宙图上的最远处是视界. 从那里看到 $t=0$ 时的情况. 在离我们"不太远"的范围内，看到的是古老程度不同的星系. 这范围之外已没有星系了. 它反映**当时**那里最早的星系尚没有形成，而不是今天没有星系. 然后我们集中到背景辐射的观测上来.

图上画出一个带阴影的薄壳. 从壳外侧"看到"的[1]宇宙气体是电离的，而从壳内侧则看到中性原子气体. 因此，壳位置 r_1 所对应的发光时间 t_1 正是宇宙气体从等离子气体向中性原子气体的转变时刻. 我们现在感兴趣的不是原子本身，而是光子. 在这短暂的转变阶段的后期，光子经过最后的一次碰撞，而变成了自由飞行的背景光子. 其中向着我们飞的，今天正好到达这里. 因此我们能用仪器俘捉它. 这就是观测背景辐射的机理. 这样看来，壳内侧的最后散射面实际是背景辐射的光源. 这光源和我们之间的空间很透明，所以是能观测的.

有了上面的分析，我们已容易理解，背景光子应当具有如下的观测性质：

(一) 如果早期宇宙是高度均匀的，那么最后散射面是球形的，球面上的气体是等温的. 这样，我们从不同方向作观测，应测到同样的辐射强度. 这也就是说，测到的背景辐射应高度各向同性.

(二) 因为背景光子在形成前很接近热平衡，所以在任一方向

① 因电离气体内光子飞行的自由程很短，我们实际上是不能观测电离气体内部的.

上观测这辐射场强度随频率的分布，它应当与理论上描述黑体辐射的普朗克公式相一致.

（三）理论上知道，背景辐射场形成时的温度约是 0.3 eV，即 3000 K. 由于宇宙的膨胀，这极遥远的光源的频率红移是 $z \approx 1000$. 不同频率的热辐射光子按同样的比例降低，等效于温度的下降. 这样，理论上估出观测到的温度应为 10^0 K.

（四）我们的银河系相对于均匀的宇宙介质必有无序运动. 观测者的运动也要引起多普勒效应，使得不同方向上测到的温度有小的差别. 考虑到银河系的速度肯定远低于光速，温度随方向的分布应当有

$$T(\theta) = T_0 + T_1 \cos\theta \tag{5.3}$$

的形式，其中的角度 θ 以银河系的运动方向为基准. 这里的 T_0 是背景辐射自身的（观测）温度. T_1 与 T_0 之比反映银河系速度 v_G 与光速 c 之比，因而应当是很微小的.

（五）早期宇宙介质应当有微小的密度起伏，即不同地方的密度应有微小的差别. 它是通过自引力不稳定性而形成星系的种子. 没有这密度起伏，后来的星系将无法产生. 密度起伏必伴随有温度起伏. 理论上估出这起伏的幅度 T_2 与 T_0 之比应超过 1×10^{-6}.

从理论上讲，背景辐射场应当具有上述这些可观测的特征. 因此，背景辐射作为热大爆炸宇宙理论的预言，它是完全能够由实测来肯定或否定的. 对我们重要的是，所有这些性质后来都得到了证实.

三、微波背景辐射的发现

背景辐射的研究历史非常具有戏剧性. 如已所述，Gamow 最初提出这理论时，学术界普遍地不信任它，甚至于没有一个人想到要用实际的观测来证实或证伪它. 到了 60 年代初，普林斯顿大学的 Dicke 和 Peebles 重新研究了这问题，并打算让两位年轻的研究

者 Roll 和 Wilkinson 制作一个仪器来测量这背景辐射的温度. 由于这热辐射的温度很低,辐射的主要部分在微波波段,且其强度也非常低,因此需要用微波辐射计做低温测量,以降低噪音. 今天看来,他们的计划当然是一个伟大的计划,而且肯定是会成功的. 可是正当这计划在进展中,他们很意外地得知,贝尔电话实验室的两位年轻人已在射电波段测到了背景辐射的信号.

贝尔实验室的 Penzias 和 Wilson 是做射电天文研究的. 他们当时的任务与天文没有关系. 它是调试一个为回声卫星计划而建造的 20 英尺角形反射天线(图 5.2),作为确定背景噪音,需要测定当天线指向天顶时的天空亮度. 天线测到的亮度通常用温度来表示,意指该温度下同频率的黑体辐射的亮度. Penzias 和 Wilson 测到的温度为 6.7 K. 他们知道其中 2.3 K 来自大气层,0.9 K 来自天线内的欧姆损耗. 于是还剩下 3.5 K 不明其来源.

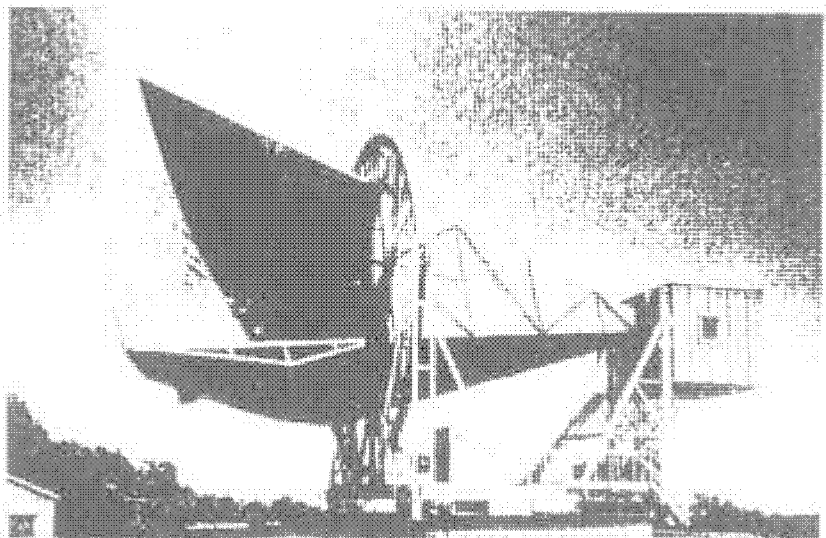


图 5.2 发现背景辐射的角形天线

天线中的不明噪音在贝尔实验室是老问题. 因实用上不一定需要弄清其来源,所以常是被忽视的. Penzias 和 Wilson 却执着地要弄清它,并花了很大的力气. 他们为此把天线拆开,发现了一对在里面做窝的鸽子. 他们清除了鸽粪,弄走了鸽子. 如此等等的努

力都没有发现这不明噪音的起因，仅排除了它来自天线内部或邻近环境的可能. 他们从而肯定了这是一个来自远处的辐射信号. 他们两人不研究宇宙学，所以完全不明白这发现的意义. 幸运的是贝尔实验室与普林斯顿大学很近，而那里的宇宙学家们太明白温度为 3K 的微波背景的意义了. 在这两组人接触讨论后，他们分头各自写了一篇论文，发表在同一期的天体物理杂志上.

Penzias 和 Wilson 的论文很短，仅一千字左右. 它只指出，在 4080 兆赫发现有一个 3.5K 的过剩天线温度；在测量的限度内，它是各向同性、非极化的和与季节变化无关的. Dicke，Peebles，Roll 和 Wilkinson 的论文则详细地解释了这信号的宇宙学意义. 这就是微波背景被发现时的历史.

微波背景辐射的发现无疑是宇宙学发展中最重大的事件之一. Penzias 和 Wilson 因此而获得了 1978 年度的诺贝尔物理奖. 但是就背景辐射问题本身而言，他们的发现是不足以令怀疑者确信的. 与上面谈到的理论预言相比，所发现的信号只符合了第一点和第三点特征. 凭单一频率（4080 兆赫）的测量，无法证实理论预言的背景辐射具有普朗克黑体辐射谱的特征（即第二点）. 因此，进一步从不同频率（波长）处去发现信号的存在，并且证实它们对应相同的辐射温度，这是确切地证实大爆炸理论的必要的一步.

四、黑体型的辐射谱

按大爆炸理论，早期宇宙气体应是高度热平衡的. 这样，观测到的背景辐射亮度随频率的分布必很接近地符合普朗克的黑体谱，即

$$I(\nu) = \frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/kT}-1}. \tag{5.4}$$

黑体谱中只含一个参量，即温度 T. 因此从任一频率作测量，即能用测到的亮度来定出辐射源的温度. 但是现在遇到的是反问题：

如何用测量来证实辐射源是高度热平衡的?要回答这个问题,我们必须在各个频率上来测量亮度,以检验这实际的亮度分布是否与黑体谱(5.4)很一致.

温度为 3 K 的热辐射的极值波长在 2 mm 左右(可提前参看图 5.3 中的理论曲线). Penzias 和 Wilson 是在 4080 兆赫,即波长为 7.35 cm 上首先发现了信号.这是热辐射的长波尾巴.进一步需要的,一是测到热辐射谱的峰,二是在短波长端测到亮度的下降.这些立刻成了研究的热点.

此后的十来年里,初步的结果是鼓舞人心的.的确,这各向同性的信号在许多波长上都存在,而且相应地定出的温度大体一致.于是越来越多的人们开始相信,Gamow 首倡的热大爆炸理论得到了证实.但是遗留问题依然存在.

要证实背景辐射具有黑体谱,重头戏在毫米和亚毫米波段的测量.在这样的波段上,地球大气的影响变得复杂而麻烦,使测量难以或不能进行.对波长短于 3 mm,须把测量仪器载在气球或火箭上来做.由于不同方法测到的结果很弥散,当时还引起过很大争议.一直到 80 年代末期,即通过 20 多年的努力,对背景辐射谱的形状还是没有共识. 1989 年发射的宇宙背景探测者(COBE)卫星在地球大气外做了测量,才使问题得到了令人满意的回答.

COBE 卫星是带着许多观测任务升天的.确定微波背景谱是它的主要任务之一.它在 30 多个波长上,每个波长用四个仪器做测量.测到的亮度分布如图 5.3 所示.小方块的大小代表测量的误差范围.图上的实线是测量结果的黑体谱拟合.两者的相符达到了惊人的程度.有这样的拟合,背景辐射的温度也在很高的精度下定了下来.它目前的值是

$$T_\gamma = 2.728 \pm 0.004\ \text{K}\quad(\text{置信度 }95\%). \tag{5.5}$$

让我们理解人们花费这么巨大的人力和物力来确定背景辐射谱的价值.

在上章开始时已指出,Gamow 的大爆炸理论其实只蕴涵着一

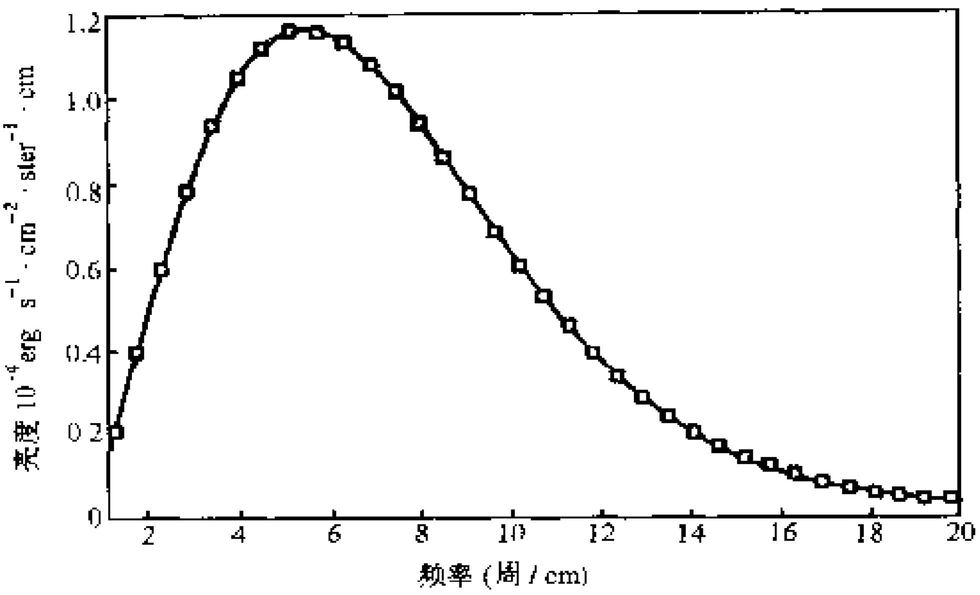


图 5.3　COBE 测到的背景光子谱

个猜测性的假定：在一切星系形成前，充满宇宙的是高度均匀的气体，且它是高度热平衡的.若能证实它是对的，那么许多其他后果都只是必然的物理推论.背景辐射的各向同性是早期宇宙均匀性的直接证据，而具有黑体型的谱则是它高度热平衡的直接证据.因此用观测来证实这两个方面，对热大爆炸理论是十分关键的.但是，在两方面都有了初步证据后，依然可以提出一个疑问：何以见得这均匀且热平衡的辐射一定是来自宇宙早期的遗迹呢？对这最后的疑问，观测到的辐射谱与黑体谱的高度吻合才是最好回答.

在星系和恒星形成后的宇宙中，各处的温度是很不均衡的.除背景场外，宇宙已没有统一的温度了.在温度不均衡的环境中实现的热平衡只能是局部的，从而也是近似的.例如太阳光球有局部热平衡，但是它实际的辐射频谱与普朗克公式差得很远.天文领域是这样，地面上的情况也一样.因此物理教科书上有一个结论：黑体辐射是一种理想化，实际上是不能实现的.

这结论无疑是对的，但是人们可能没有意识到，十分地接近理想黑体的辐射是存在的，那就是在早期宇宙中.这或许是惟一的可能！因为只有在那里，全空间的物质都处于同样的物理条件下，这

才能达到高度的热平衡. 于是,其光子频谱必与普朗克公式十分接近. 在已有了星系的宇宙中,这确是无法实现的. 在背景辐射被发现后,大爆炸理论的反对者曾企图设计另一种机制来解释它,但是它们始终没有成功. 困难就在于他们无法使其辐射谱与黑体谱有高度的符合. 因此,对背景辐射谱作精确测定的价值是把热大爆炸理论的正确性肯定到了无法争议的地步.

五、银河系的绝对运动

在前面所画的观测宇宙图中,假定了处于中心位置的银河系是静止的. 对河外星系的观测表明,它们除了随宇宙膨胀外还有自身的运动. 它就是天文学的本动. 银河系没有特殊性,它自然也应有本动. 我们实际上是在运动着的物体上观测背景辐射. 这样就应考虑观测者的运动所引起的多普勒效应.

在垂直运动方向上没有多普勒效应,测到的背景辐射温度(记为 T_0)应与静观测者一致. 在我们运动的前方,频率红移减小了一些,相当于温度高了一些. 同理,背着的方向上测到的温度低了一些. 具体地导出温度随方向的变化,就是前面提到过的结果:

$$T(\theta) = T_0 + T_1\cos\theta, \tag{5.6}$$

这种方向特性与偶极辐射一样,所以叫偶极各向异性. 变化幅度 T_1 与观测者速率 v 的关系为

$$T_1/T_0 = v/c, \tag{5.7}$$

这里已注意到 $v \ll c$. 一旦测到这样的各向异性,就可以定出银河系的本动速度.

从 1978 年开始,用单一频率作测量,很多人发现了这种与 $\cos\theta$ 成正比的温度变化,表明上面的理论预言是对的. 变化幅度用 1996 年的新结果是

$$T_1 = 3.353 \pm 0.024\ \mathrm{mK}, \tag{5.8}$$

即 $v \approx 10^{-3}c$. 扣除地球绕太阳的运动,推知太阳的运动速率为

360 km/s，方向指向银经 $l=264.3°$，银纬 $b=48.2°$. 天文学家早已测定过太阳绕银河系中心的速度. 再扣除这部分运动，进一步推出银河系的速度约为 600 km/s.

人们常把这速度叫银河系的绝对速度. 提起绝对速度会使我们联想起一百年前的迈克耳孙-莫雷实验. 它的目的是测量地球的绝对速度(当时指对静以太的速度). 测到的结果为零，使爱因斯坦意识到以太不是物理实体，从而从新确立了相对性原理，建立了狭义相对论. 那么现在测到银河系的绝对速度意味着什么?

确实，均匀地弥漫于全宇宙的背景光子的存在，为我们提供了一个能定义观测者绝对静止的物理基准：若他发现背景辐射温度没有偶极各向异性，则他是绝对静止的. 上面谈到的银河系的运动就是相对这样的静止系而言的，因此可以称它为绝对运动. 那么，这是否意味着牛顿绝对空间概念的复活?

考虑到宇宙介质和它占据的空间都在膨胀，全局性的参考系是不存在的，因为参考系须由一群相对静止的物体组成. 实际使用的参考系都只是局域的测量基准. 上面讲的静止系也只是局域的. 以观测到背景辐射为各向同性作准则，不同的地方有各自的静止系. 它们之间是随宇宙的膨胀而有相对运动的. 所以，这样的静止系与牛顿的绝对空间很不一样.

无论如何，有了上面的准则，动与不动就不完全是相对的了. 在任一地点可有无数个相对地运动着的惯性系. 但是其中确有一个是特殊的. 在它看来，背景辐射是没有偶极各向异性的，而从别的看来都有. 从这意义上讲，绝对空间的概念得到了一定程度上的恢复. 但是这与狭义相对论并没有直接的冲突，更不意味着对狭义相对论的否定.

六、原初扰动的痕迹

背景辐射证实了早期宇宙很均匀，那么后来星系的形成只能

靠自引力的不稳性. 于是均匀的背景上必须有原初密度扰动. 尽管今天还不很清楚原初扰动的产生机理①,但是只要它产生在背景辐射形成之前,就必定应在背景辐射上留下印记. 密度的不均匀可通过温度的不均匀反映出来,从而是能观测到的.

上节讨论的偶极不均匀是观测者运动的多普勒效应,而不来自最后散射面上的实际起伏. 这样,背景辐射的温度分布还应有四极和更多极的各向异性. 它们才是原初扰动的痕迹.

因为这些小扰动的增长很慢,要今天已形成大量的星系,当时的扰动幅度不能太小. 体现的温度起伏上,至少要 $\delta T/T_0>1\times10^{-6}$,即至多比偶极起伏小三个量级. 偶极起伏是 mK 的量级,它应是 μK 量级. 这么小的温度各向异性很难测量.

从 70 年代末发现温度的偶极各向异性后,人们就开始了寻找它的努力. 由于受测量精度的限制,很长时间里得到的都是零结果,从而给出扰动大小的上限. 起初肯定了 $\delta T/T_0<10^{-4}$. 到 80 年代末,随测量精度的提高,扰动上限已降低到了 $\delta T/T_0<1\times10^{-5}$. 这使研究星系形成的宇宙学家感到了困惑. 如果精度再提高一个量级,而得到的仍是零结果,将意味着 $\delta T/T_0<1\times10^{-6}$,那么星系就来不及产生了.

在这样的关键时刻,COBE 卫星又一次做出了重要的贡献. 1992 年,COBE 的另一套测量仪器 DMR 终于测到了人们期望看到的温度起伏. 它实际测量的是不同方向上相隔 10°两点的温度差的平均. 得到的结果是

$$\delta T/T_0 = 5\times10^{-6}. \qquad (5.9)$$

这个值大体上落在人们预料的范围之内. 至此,背景辐射上应看到的性质都看到了. 作为对大爆炸理论的检验,这应当说是交出了一份完美的答卷.

尽管背景辐射的观测已十分成功,但新的研究计划仍在提出.

---

① 随机的热涨落肯定有,但它作为原初扰动是不够的.

当然动机已经转移.现在的目的是要从中攫取更多关于宇宙学的信息.

在1998年以来的两年里,有过两个气球观测计划.一个叫BOOMERanG,另一个叫Maxima.它们的目的是测量微波背景上的温度扰动谱.借助测到的扰动谱能推断出宇宙的整体参量.图5.4在Ω_λ-Ω_m平面上示出了与BOOMERanG的观测结果相拟合而定出的Ω_λ和Ω_m值.椭圆状的实线内是它们在95%置信度上的取值范围.这窄长的范围强烈地表明了$\Omega_\lambda+\Omega_m=1$.它证实了理论家十年来的一个猜想,所以是很重要的信息.图上椭圆状的虚线是Perlmutter组拟合红移-距离关系得到的结果(参看图3.3).这两个结果的结合,得到了$\Omega_\lambda=0.65$左右.它的意义已在第三章中充分讨论了.人们专门去细致地测量这扰动谱,是想把它用作研究星系形成的初条件,因此我们在最后一章中还会谈到它.

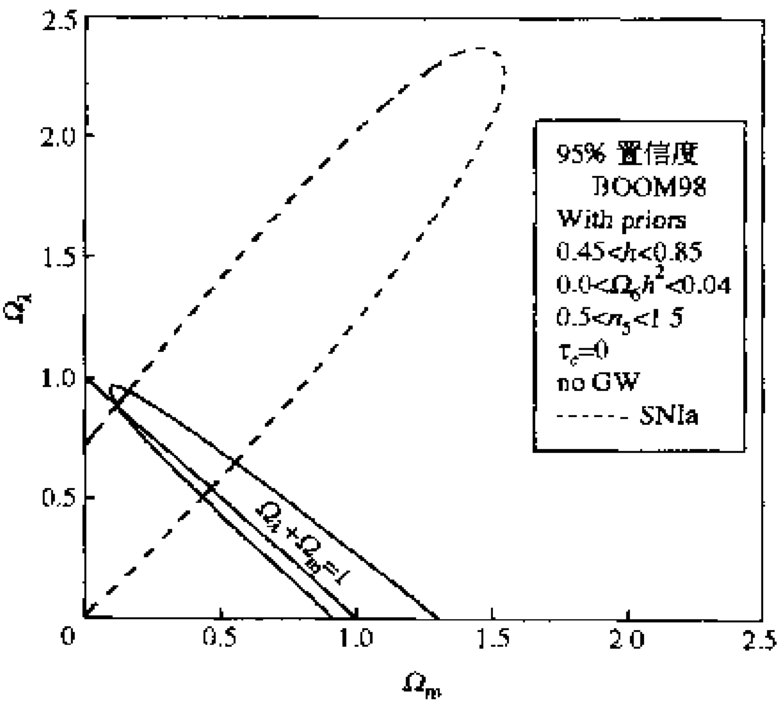


图5.4 BOOMERanG测量的拟合结果

第六章　原初的核合成

背景辐射理论的成功，证实了宇宙确曾经很热．至少它曾热到过 10^4 K．那么更早期宇宙温度曾高到化学元素也无法存在是必然的推论．它们只能是后来合成的．理论处理后知，宇宙中第一批元素的是氢(^{1}H)、氘(^{2}D)、氚(^{3}T)、氦(^{3}He 及 ^{4}He)和锂(^{7}Li)．它们的合成阶段发生在宇宙年龄为 3 分至 1 小时．天文测量已初步证实了理论上算出的产额．从而说明，前推到宇宙年龄仅为几分钟，大爆炸理论算出的结果依然是符合事实的．

一、最初的原子核合成过程

当宇宙年龄为 1 秒时，介质温度为 1 MeV，即 10^{10} K．当时对介质密度有主要贡献的组分粒子是静质量小于 1 MeV 的一切粒子，即光子、正反中微子和正反电子．正反核子(质子和中子)在温度超过 1 GeV 时能热碰撞产生，并大量地存在．后来它们中绝大部分都成对地湮灭了．残存核子的数量已很少，成了宇宙中的次要组分．但是对于现在要讨论的原子核合成问题讲，这些核子却是问题中的主角．

既然气体中有质子也有中子，它们相碰就能结合成最简单的原子核——氘(^{2}D)，

$$p + n \longleftrightarrow {}^2D + \gamma \qquad (6.1)$$

氘的结合能是 2.2 MeV. 因此，能量超过 2.2 MeV 的高能光子可以把氘核瓦解，即上述过程是可逆的. 当 $T=1$ MeV，足以使氘瓦解的光子数目依然太多. 在正反过程的统计平衡下，氘的数量微不足道. 真正的氘合成须等待宇宙温度的下降，以使氘被瓦解的几率变得次要，甚至可忽略.

要计算氘合成的开始时刻，需要知道核子数与光子数之比 η. 这是原初核合成理论的重要参量，但是我们只知道它的值在 10^{-10} 到 10^{-9}之间. 正因为每一个核子周围平均有 10^9 到 10^{10}个光子，所以要等到宇宙温度降到 0.1 MeV 左右，能量大于 2.2 MeV 的高能光子才足够少，以致氘核的光分裂实际上失效. 这被看作是氘核合成的开始. 相应的时间约是 $t=3$ min. 人们常通俗地讲，到年龄为 3 分钟时宇宙中才开始有最早的原子核，其含义即在于此. 认真些讲，氘合成的开始时间与 η 的大小有关. η 大反映光子数相对地少，因此氘的光分裂失效早，即氘合成发生得早.

让我们先作定性分析. 一旦氘合成开始，后继的连锁反应就能接着发生了. 算过后知道，实际起主要作用的过程有

$$^2\mathrm{D} + \mathrm{p} \rightarrow {}^3\mathrm{He} + \gamma, \tag{6.2}$$

$$^2\mathrm{D} + {}^2\mathrm{D} \rightarrow {}^3\mathrm{He} + \mathrm{n}, \tag{6.3}$$

$$^2\mathrm{D} + {}^2\mathrm{D} \rightarrow {}^3\mathrm{T} + \mathrm{p}, \tag{6.4}$$

$$^3\mathrm{He} + \mathrm{n} \rightarrow {}^3\mathrm{T} + \mathrm{p}, \tag{6.5}$$

$$^3\mathrm{T} + {}^2\mathrm{D} \rightarrow {}^4\mathrm{He} + \mathrm{n}, \tag{6.6}$$

$$^3\mathrm{He} + {}^2\mathrm{D} \rightarrow {}^4\mathrm{He} + \mathrm{p}. \tag{6.7}$$

图 6.1 画出了相应的反应网络. 在产生氘之后，主要合成的核是 ^{3}T、^{3}He和 ^{4}He，即原子序数为 1 和 2，原子量为 2，3 和 4 的轻原子核.

为什么热核反应链在原子量 A=5 的地方中断了呢?

从核物理事实知道，对原子量 A 从 1 至 200 多，几乎任一 A 值都有相应的稳定或亚稳的原子核存在，但正好有两个空缺：A 为 5 和 8 是没有的. 原初核合成过程每次只产生原子量增加 1 的

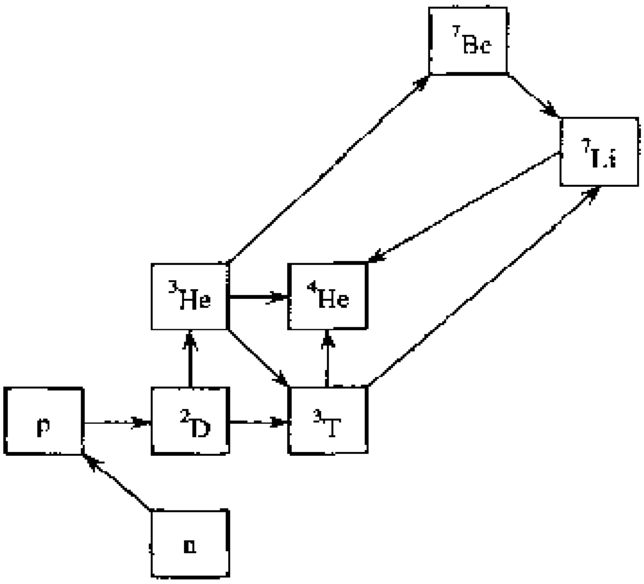


图 6.1 原初核合成的反应网络

核. 所以由于 A＝5 没有稳定核，就使得反应链断了. 早期宇宙核合成最终主要是产生大量的 ^{4}He.

在反应阶段的后期，有可能出现原子量增加更多的反应. 例如随着 ^{4}He的积累，两个 ^{4}He核能碰撞产生原子量为 8 的铍核. 可惜原子量为 8 的锂、铍、硼等核也都极不稳定，所以反应链还是没有延续下去. 细算起来，最后还会合成出少量的 ^{7}Li 和 ^{7}Be(示于图 6.1 的右上方). ^{7}Li 的产生方式主要有两种，

$$^{4}He + {}^{3}T \rightarrow {}^{7}Li + \gamma \tag{6.8}$$

或
$$^{4}He + {}^{3}He \rightarrow {}^{7}Be + \gamma \rightarrow {}^{7}Li + e^{-} + \bar{\nu}_e. \tag{6.9}$$

所产生的^{7}Li 还会与 p 相碰而部分转回^{4}He，即

$$^{7}Li + p \rightarrow {}^{4}He + {}^{4}He. \tag{6.10}$$

因此，净效果仅是产生微量的 ^{7}Li. 铍的净产额更低，我们将不讨论它. 归根到底，原初核合成只产生了周期表的前三种元素及其同位素. 第四和第五号元素也产生了一点点，但太少了. 这过程完全不能产生第六号元素碳，或序数更大的元素. 今天的碳、氮、氧和更重的元素都是恒星制造出来的.

热核反应需要反应前的原子核有足够的热动能，以克服核与

核之间的库仑排斥力，使它们靠近到核力能起作用的程度．温度必须很高的理由就在这里．但是宇宙气体却是在随着膨胀而降温．约1小时后，气体的温度降到了 0.02 MeV（即约 10^8 K）．这时在核与核的热碰撞中，核力已无法起作用．于是一切热核反应都停止，原初的核合成过程也就结束了．在宇宙早期演化中，原子核的形成只是穿插在其中的一个短阶段．粗略地讲，这阶段的时间是从3分到1小时，宇宙温度是从 10^8 K 到 10^9 K．

在原初核合成结束后，宇宙介质变成了等离子状态的气体．其化学组分主要是氢和氦．氢就是核合成结束后所剩余的质子，而氦（^{4}He）是合成的主要产物．以质量衡量，氢约占 3/4，氦约占 1/4．其他原子核占万分之一以下．若先不考虑那些微量元素，我们对 ^{4}He 的产额可做简单的估算．

设核合成开始时质子和中子的数密度为 n_p 和 n_n（下一节将论证 $n_p>n_n$）．注意到合成一个氦核消耗2个质子和2个中子，所以在全部中子耗尽后尚有余下的质子．它们就是后来的氢核．氦的丰度 Y_4 定义为它在气体中的质量百分比，即

$$Y_4 = \frac{n_4}{n_1 + 4n_4}, \tag{6.11}$$

其中 n_1 和 n_4 分别是氢和氦的数密度．因为 $n_4=n_n/2$ 以及 $n_1=n_p-n_n$，上式可化成

$$Y_4 = \frac{2n_n}{n_n + n_p} = \frac{2}{1 + (n_n/n_p)^{-1}}. \tag{6.12}$$

从这里看出，氦的丰度取决于核合成开始时的中子和质子数之比．

Gamow 最初估出，$n_n : n_p = 1 : 7$，并由此推断出大爆炸初期合成了 1/4 的氦．当时已有太阳系氦丰度的实际测量，Y_4 约比 1/4 略多．大家都知道恒星在把氢烧成氦，但是远不可能产生那么多．Gamow 就用原初核合成解释了这氦丰度的来源．但是在学术界还根本不相信膨胀宇宙理论时，这种成功是不足以改变人们的态度的．对原初核合成问题的认真理论研究和实测检验是背景辐射被

发现后的事.

二、影响产额的两个因素

背景辐射的存在,表明星系形成前的宇宙确是均匀的高温气体.至少宇宙温度曾高达 10^4 K.那么,在更早的阶段曾有过更高的温度,并发生过原初的核合成,这是一个必然的物理推论.正是基于这样的信心,在背景辐射发现后,人们立即对这问题进行了很详尽的理论研究.

从核物理的角度看,这种温度条件下的核反应规律是十分清楚的.因此人们的目标不再是作估算,而是用核反应的理论方法来认真算出原初核合成的产额,以便得到定量的实测检验.当然,我们在这里还是宁肯作一些半定量的分析.上节已说明,中子与质子的粒子数比对产额有决定性的影响,因此我们从这问题开始讨论.

在 $T>1$ MeV 时,气体中的质子和中子是能通过弱作用而相互转化的,即

$$p + e^- \longleftrightarrow n + \nu_e, \tag{6.13}$$

$$p + \bar{\nu}_e \longleftrightarrow n + e^+. \tag{6.14}$$

正反过程都频繁地发生,使得质子和中子的数密度都满足玻尔兹曼分布律,因而它们之比为

$$n_n/n_p = e^{-\Delta m/T}, \tag{6.15}$$

其中 $\Delta m = m_n - m_p = 1.29$ MeV 是中子和质子的质量差.由此首先看出,因为中子比质子重,所以中子一定比质子少.同时它还告诉我们,中子与质子数之比是随温度下降的.

要质子与中子保持这样的转化,这种弱过程的发生率必须远超过宇宙的膨胀率,即 $\Gamma_{弱} \gg H$.一旦 $\Gamma_{弱}$ 和 H 的大小可比拟,弱作用将退耦,即中子与质子不再能这样转化了.于是,中子和质子的数目比冻结了下来.这种冻结开始在温度 T_f 略小于 1 MeV 的时刻.冻结后的比例约为 1∶7.由于核合成开始时的中子-质子比

很重要，因此冻结温度T_f的确切大小是决定最后产额的一个关键量.

在 n_n/n_p 冻结后，中子还会通过衰变而转化为质子. 这衰变的半衰期是 10 min，而当时宇宙年龄还只有几秒. 衰变实际还来不及发生. 我们知道，在 $T\approx1$ MeV 时，氘的合成也还没有开始. 上节已讨论过，氘合成的开始由其反过程——氘的光分裂失效来标志. 它的温度约在 $T_D=0.1$ MeV 上下. 相应时刻是 $t=3$ min 前后. 所以在氘合成开始时，中子已部分地衰变了. n_n/n_p 因而减小了一些. 这样看来，氘合成的开始时刻 t_D(温度 T_D)又是一个影响产额的关键量. t_D 越小，则 n_n/n_p 越大，最终合成的 ^{4}He也相应地越多.

那么 T_f 和 T_D 的确切大小又分别取决于什么?上节已讨论过，T_D 的大小取决于气体中核子与光子的数目比 $\eta(=n_N/n_\gamma)$. η 越大则 T_D 越大(即 t_D 越小). 现在来论证，T_f 的大小与自然界中微子有多少种(记作 N_ν)有密切关系.

弱作用的退耦是其发生率 $\Gamma_{弱}$ 与宇宙膨胀率 H 相竞争的结果. 从宇宙膨胀的动力学方程知，

$$H^2=(\dot{R}/R)^2=(8\pi G/3)\rho. \qquad (6.16)$$

每一种质量小于 1 MeV 的中微子都对当时的宇宙密度有一份不可忽略的贡献. 因此，若中微子的种数 N_ν 越大，当时宇宙的膨胀率 H 也越大. 这样，弱作用的失效就发生得越早，于是参与核合成的中子数将相对较多. 这就是自然界中微子的种数会影响原初核合成产额的道理. 它实际上影响的是当时宇宙膨胀的快慢.

在对产额作理论计算中，当然会涉及许多核反应截面的数据. 但是这些数据都已在核实验室中测量过了，因此这不给产额计算带来困难. 计算需要而我们不确切知道的参量只有两个，那就是 N_ν 和 η. 于是人们只能把它们当输入参量来处理.

图 6.2 画的是计算出来的元素丰度或数密度的对数随温度(即时间)的变化. 那里已输入了 $N_\nu=3$ 和 $\eta=3\times10^{-10}$. ^{4}He含量用

其丰度 Y_4 标志，而其他核的含量都用它与氢的粒子数比 y 来标志. 例如氘用 $y_2=n_D/n_H$ 刻画. 这样描写仅是为了便于与实测比较，而没有理论上的原因. 从算出的曲线看，每种核的量是连续变化的，并没有严格的"开始". 我们作的是半定量分析，才把氘含量达到一定程度时叫核合成的开始. 从这些曲线也具体地看到，氘合成在先，紧接着是氚、^{3}He 和 ^{4}He 含量的迅速增长. 锂和铍的合成是在 ^{4}He 已很多时才开始的.

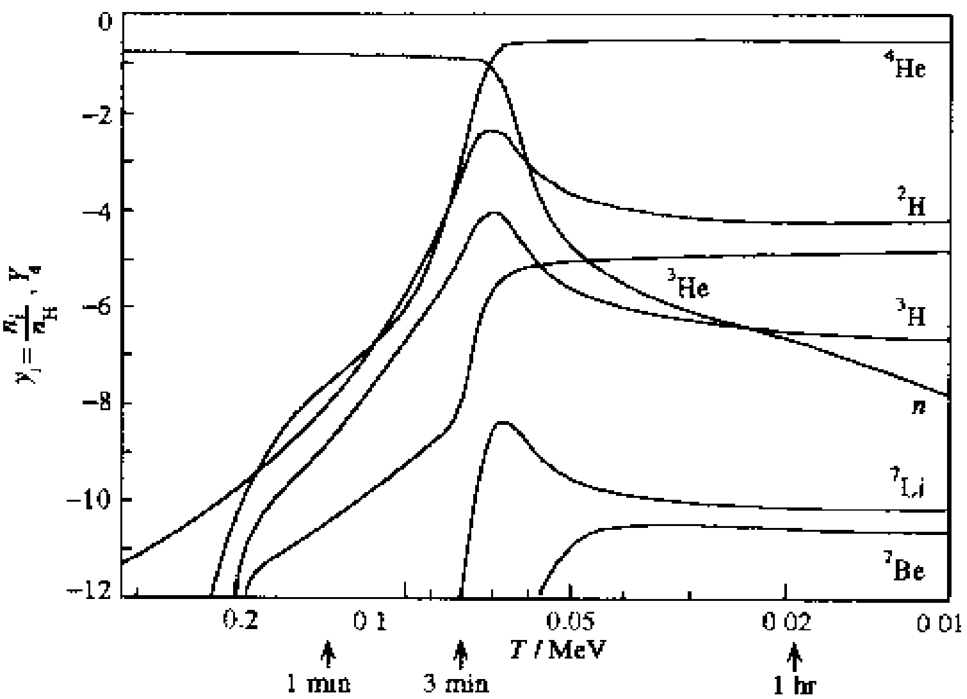


图 6.2　核合成的发展过程

图上也画出了中子数的减少过程. 若以中子的显著减少为核合成开始的迹象，那么大约这时刻是 $t=3$ min，相应温度约是 0.1 MeV. 到核合成进行了约 1 小时，宇宙温度下降到 0.01 MeV 左右，各种核的含量不再有显著变化. 这是核合成阶段结束的标志. 图上清楚示出，过程结束后留下的不同核的含量有很多量级的差别.

我们真正感兴趣的是最终产额. 图 6.3 画出了 ^{2}D, ^{3}He, ^{4}He, ^{7}Li 的产额作为 η 的函数，中微子的种数 N_ν 被取定为 3. 原初的氘

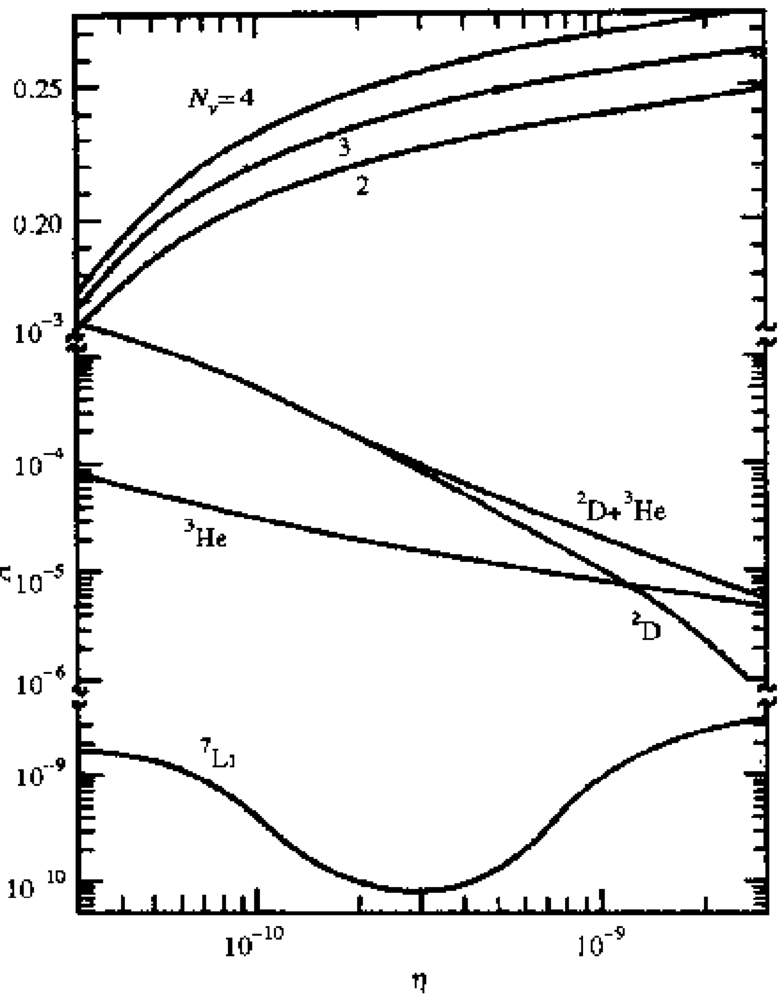


图 6.3 核合成产额随 η 的变化

后来会衰变成 ^{3}He，因此其产额已计入^{3}He 的产额 y_3 中了. y_2（氘的产额）和 y_3 都是 η 的降函数. 它们的大小在 10^{-5}到 10^{-4}之间. 图上多画了一条 y_2+y_3 的曲线是后面与实测比较的需要. Y_4 是 η 的升函数，其定性原因已讨论过. N_ν 对 Y_4 的影响将在下一节中讨论. 锂产额 y_7 在我们感兴趣的范围内是 10^{-10}至 10^{-9}，它不是 η 的单调函数，而存在着一个"谷".

三、^{4}He 原初丰度的实测推断

原初核合成理论提供了很可靠的产额结果，进一步需要的是

对它的实测检验. 让我们注意,理论上算出的是核合成结束时的气体组分,当时宇宙年龄仅为 1 小时. 今天测量的却是已演化了 150 亿年的宇宙中的近期组分. 在这两个时间之间曾发生过许多恒星的形成和死亡. 我们知道,恒星的核燃烧会产生氦、碳、氮、氧等化学元素,而在死亡时又会部分地把它合成的元素抛到外部(星际)气体中来. 所以恒星演化是改变了气体的化学组分的. 意识到这一点,在把理论结果与实测比较前,须先在实测值中扣除恒星过程的影响. 这可是一个十分困难的任务. 幸好至少对 ^{4}He,人们已有了一种合理的方法.

天文学家把测量对象集中在河外 HII 区. 它们是主要由电离氢和氦组成的气体云. 许多 HII 区的氧、氮和碳的含量比太阳系低很多. 氧、氮和碳都是恒星过程的产物. 它们的含量低,表明这些天体形成得早,因而受恒星过程的影响小. 这样,其中的 ^{4}He 丰度应较接近原初值. 可惜从许多不同 HII 区中测到的 Y_4 依然很弥散. HII 区中的碳、氮、氧含量也能测定. 从道理上讲,原初核合成不产生这些核,因此它们的含量(y_C, y_N 和 y_O)是恒星影响程度的标志. 作统计分析也能看出,碳、氮或氧含量越小的 HII 区中 ^{4}He 也越少. 利用这道理,人们把不同 HII 区的 Y_4 看成 y_O(氧较好测定)的线性函数,来与实测结果相拟合. 然后内插到 $y_O=0$, 以消除恒星的影响. 这样得到的结果记作 Y_{4p},可理解为由实测推断出的 ^{4}He 原初丰度. 它应当与理论产额相一致.

图 6.4 取用的是最早(1991 年)这样做时所用的数据图. 线性拟合得到

$$Y_4 = (0.229 \pm 0.004) + (1.3 \pm 0.3) \times 10^2 y_O. \quad (6.17)$$

内插结果, 定出 $Y_{4p}=0.229\pm0.004$. 从图上就能直观看到,这样做的可靠度会较差. 同一组研究者又用 y_N 或 y_C,按同样方法对 Y_4 做了关联分析. 这样又能独立地定出 Y_{4p}的值. 把三个结果结合在一起,他们发现最后只能定出一个两位有效数字的范围:

$$Y_{4p} = 0.22—0.24. \quad (6.18)$$

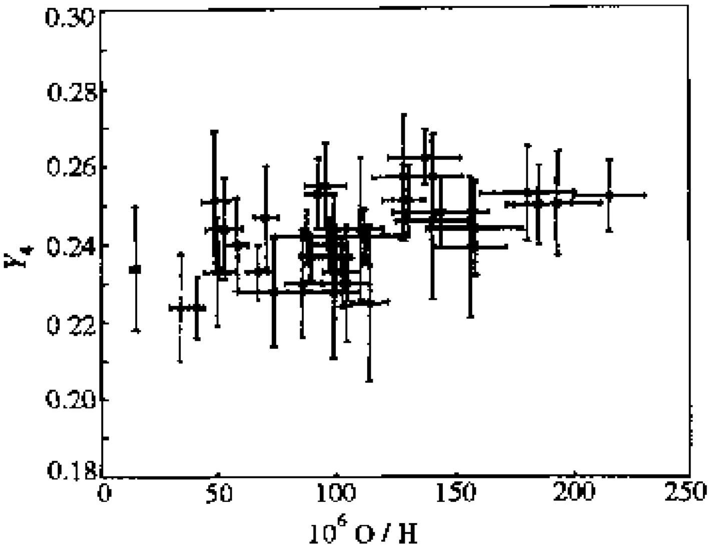


图 6.4 ^{4}He 与 O 的丰度关联

到 1994 年，人们想把 Y_{4p}的第二位小数可靠地确定下来. 在对内插手续做了改进后，得到的结果是

$$Y_{4p} = 0.232 \pm 0.003. \quad (6.19)$$

这里的误差仅是统计的. 此外还有若干会导致系统误差的因素. 估计系统误差的大小是 0.005 之内. 这样，采用 2σ 的统计误差(95%置信度)，再计及系统误差，Y_{4p}的范围是

$$Y_{4p} = 0.221—0.243. \quad (6.20)$$

实际上第二位有效数字仍然没有确定下来. 无论如何，我们对它有了一个较可信的范围.

顺便指出，这结果表明，今天宇宙中的 ^{4}He 确实大部分来自宇宙早期，而恒星产生的仅是次要部分.

四、中微子的种数问题

1972 年，粒子物理实验发现了质量约为 2 GeV 的 τ 轻子. 人们才意识到中微子有三种(常被称为有三代)：与电子($m=0.511$ MeV)同代的中微子 ν_e，与 μ 子($m=106$ MeV)同代的中微子 ν_μ 和与 τ 子($m=1784$ MeV)同代的中微子 ν_τ. 那么随着加速器能量的

提高，是否还会有更多代轻子被发现出来？当时的理论家对此还没有想法. 若用其他粒子实验数据来找约束，当时能用的只有 K^0 介子的衰变. 对中微子种数得出的限制是 $N_\nu<10^5$. 这几乎等于没有提供约束.

70 年代中期，原初核合成的研究初步给出了一个很苛刻的限制：$N_\nu\leqslant4$. 这当然引起了粒子物理学家的注意. 问题在于粒子物理是很成熟的学科，而宇宙学当时尚在襁褓时期，由宇宙学向粒子物理提供的信息可信吗？这是一个合理的疑问.

为了简单，让我们用今天的方式来解释宇宙学的预言是怎么做出的.

图 6.5 中单独地画出了 $N_\nu=3,4$ 时 Y_4 随 η 变化的理论结

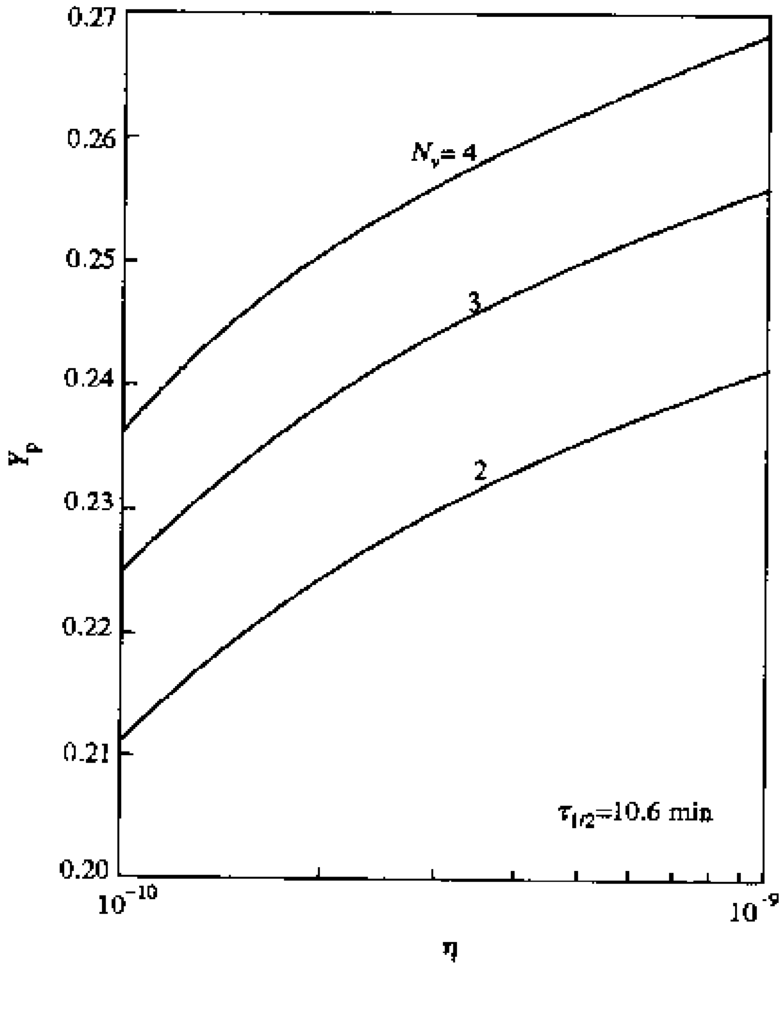


图 6.5 N_ν 对 ^{4}He 产额的影响

果. 这实际上是图 6.3 的上部的细化. 计算结果表明，每多一代中微子，^4He 的产额将约增大 0.014. 因此，若 N_ν 很大，^4He 的理论产额会变得太大，从而与事实显然不符①. 当时的论据实际上就是这样. 只因为当时 Y_4 的值尚没有这么确定，论证要复杂一些. 按今天的数据，Y_4 的实测推断值范围在 0.221 到 0.243 之间. 由这些曲线已易于看出，$N_\nu=3$ 符合很好，$N_\nu=4$ 不排除，而 $N_\nu=5$ 则是不行的. 这就是宇宙学的论据.

此后宇宙学在发展，粒子物理也在进展. 宇宙学家愈来愈相信关于 $N_\nu \leqslant 4$ 的预言很可靠. 粒子理论也在论证轻子总共只有三代. 但是这问题上最权威的结论只能来自粒子物理实验. 1985 年 5 月，正负电子对撞机的实验定出了 $N_\nu<14$，同年同月的正反质子对撞机实验又把限制缩小到 $N_\nu<5$. 它与宇宙学提供的限制一样了. 这当然对襁褓期的宇宙学是极大的支持. 到 1989 年，新一代加速器(LEP)产生了上百万个 Z^0 粒子. 实验家由它的衰变分支比确切地定出了 $N_\nu=2.96\pm0.06$. 这对粒子物理肯定是一个重要结果. 从我们的角度讲，它也是原初核合成理论的正确性的强烈证据.

五、其他轻核的原初丰度

对于氘、^3He 和 ^7Li，我们也希望能得到它们原初丰度的实测推断值. 但是尚没有很可靠的结果. 无论如何，它们对于全面地检验理论十分重要，我们也作些简单讨论.

先讨论氘. 人们肯定地弄清了，一方面是恒星在形成前夕(即主序前阶段)必先把其中的氘燃烧成了 ^3He，另一方面是任何天文环境都没有可能显著地产生氘. 这样看来，宇宙中的氘应是随着恒星一代代的形成而单调地减少的. 从实测讲，由于氘的丰度太低，

① 意指即使取 $\eta=1\times10^{-10}$，相应的理论产额也超过了实测允许的上限.

天文学家只能在太阳系内，或邻近的星际气体中测量氘与氢(或氦)的粒子数比. 按上面的道理，这种实测值 y_2 能为原初值 y_{2p} 提供了一个可靠的下限. 由此得到的具体结果是

$$y_{2p} > y_2 \approx 1.8 \times 10^{-5}, \qquad (6.21)$$

这虽然只是一个对 y_{2p} 的约束，但应是很可靠的.

1994 年以来，有若干天文观测组从类星体前方的吸收云中测到了氘的吸收线，并由此推出了吸收云中的 y_2. 这些吸收云的红移都很高，说明它们是很古老的天体. 其中的氘含量应当与原初值 y_{2p} 很接近，因此测到这些数据对检验原初核合成理论十分重要. 可惜的是，不同古老天体中的氘含量不应有很大差别，而从不同吸收云中测定的 y_2 值很弥散. 已发布的最大值是 $y_2=(1.9-2.5)\times 10^{-4}$，而最小值是 2.4×10^{-5}. 大与小差了一个量级. 因此原初氘丰度究竟多大的问题依然没有解决.

关于推断 ^{3}He 的原初丰度，事情更复杂. 不同质量的恒星的影响不一样. 有的在演化中会消耗 ^{3}He，有的会产生 ^{3}He. 作为综合效果，宇宙中的平均 ^{3}He 含量并不单调地变化. 在实测方面，因为 ^{3}He 丰度较低，像氘一样，它也只有近处的测量. 要把它和原初值联系起来，需要关于星系中恒星形成和演化的许多知识，而这些知识是人们尚不具备的. 因此至今所做的尝试，都蕴涵有不知是否正确的假设. 让我们讨论一种常被引用的模型.

假设银河系中所有的恒星是同时在原初气体中形成的，再假设恒星死亡时向星际抛出的化学元素会很快在全星系中均匀化. 那么，在对氘和 ^{3}He 做综合考虑后，就可以导出如下的不等式：

$$y_2 + y_3/g > y_{2p} + y_{3p} \equiv y_{23p}, \qquad (6.22)$$

其中 y_2 和 y_3 是实测量，加下标 p 代表相应的原初量. 这不等式中涉及了星系演化中的 ^{3}He 残存率 g. 它被定义为现在已死亡恒星抛出的 ^{3}He 量与恒星形成时的 ^{3}He 量之比. 天体物理学家对恒星的演化比较了解，因此能理论上估出 g 的上限. 合理地估计，有 $g>1/2$，而取用 $g>1/4$ 则是安全的. 这样，利用实测的 y_2 和 y_3，这

不等式就能为我们提供原初值 y_{23p}的上限. 具体地讲，由此定出的是

$$y_{23p} < 1.0 \times 10^{-4}. \qquad (6.23)$$

这分析方法和相应的结果在研究文献中常被用到. 我们想强调，由此得到的公式(6.22)到底可靠到什么程度其实是不大清楚的. 关于 y_{23p}(或 y_{3p})，目前还没有很可靠地从实测推出的方法.

最后简单讨论一下 ^7Li 原初丰度的实测推断. 20 世纪 80 年代前，从陨石、邻近星际气体或星族 I(较晚近形成的)恒星观测，定出的 y_7 值在 10^{-9}左右(差因子 2). 人们知道宇宙线散裂能产生 ^{7}Li，此外它也容易在 $T>2\times10^{-6}$ K 的环境中被烧掉(例如太阳大气中 y_7 仅有 10^{-11}). 因此几乎看不到从这些观测结果中推断原初丰度的可能.

1982 年开始有人从重元素含量比太阳低一个量级以上的晕星中测到了 ^{7}Li. 重元素含量低暗示了它们形成得很早，因此这些恒星大气中的 ^{7}Li 量可能与其原初值较接近. 在分析从这类恒星的 y_7 与恒星表面温度 T 的关联后，发现对表面温度低于 5500 K 的恒星，y_7 随恒星温度下降而下降；但是 $T>5500$ K 的恒星却有接近相同的 y_7，从而在关联图上形成了一个高台(图 6.6). 相应的 ^{7}Li 与氢的粒子数比为

$$y_7 = (1.1 \pm 0.4) \times 10^{-10}. \qquad (7.24)$$

后来有论据说明这些恒星大气中的 ^{7}Li 就是原初 y_{7p}的好的近似，但是争议很大. 对于论据很细腻而结论又不十分明确的地方，我们就不作过多的讨论了.

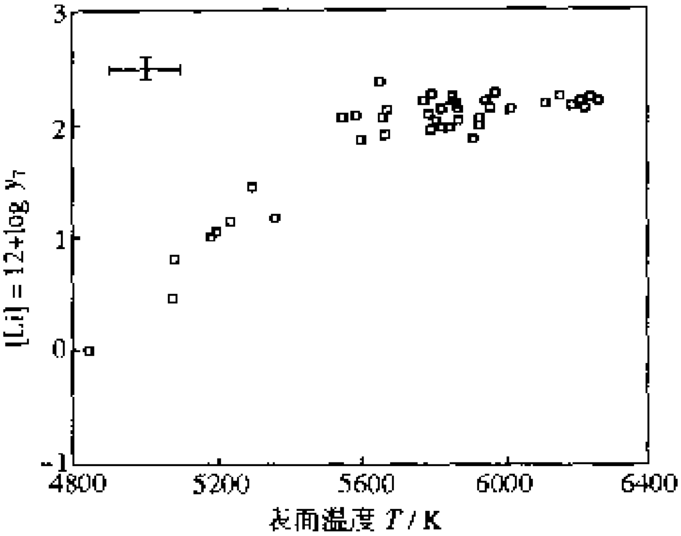


图 6.6 晕星中锂的含量

六、理论的全面检验

1984 年 Yang 等人①第一次尝试用实测丰度对原初核合成理论作全面的检验. 检验的思想是这样的. 分别用实测推断的原初值 Y_{4p}, y_{2p}, y_{23p}和 y_{7p}与理论产额比较(参看理论产额曲线),从而独立地得到 η 应取值的区间. 若这些区间没有公共的范围,则是对理论的否定. 反之,若存在一个公共的范围,则表明理论得到了实测的支持. 相应定出全面适合的 η 范围就算是原初核合成理论对核子数密度的预言.

1984 年的实测推断值尚没有这么确定. 让我们引用这研究组在 1991 年的结果②. 他们先用 $y_{2p}>1.8\times10^{-5}$和 $y_{23p}<1.0\times10^{-4}$,定出

$$2.8\times10^{-10}\leqslant\eta\leqslant6.8\times10^{-10}. \qquad (6.25)$$

① Yang Turner Steigman Schramm and Olive, Astrophysical Journal 281(1984) 493.

② Walker Steigman Schramm Olive and Kang, Astrophysical Journal 376(1991) 51.

再用晕星中的 y_7 值($\log y_7 \leqslant -9.85$),加上锂产生的修正,定出

$$1.6 \times 10^{-10} \leqslant \eta \leqslant 4.0 \times 10^{-10}. \quad (6.26)$$

这两个区间有一个共同的范围为

$$2.8 \times 10^{-10} \leqslant \eta \leqslant 4.0 \times 10^{-10}, \quad (6.27)$$

被他们认为是 η 的实际值. 取 $N_\nu=3$,这 η 范围对应的 ^{4}He 丰度在 0.236 到 0.243 之间. 它与由观测推断的原初丰度 $Y_{4p}=0.23\pm 0.01$ 相符. 由此他们宣称,理论结果与所有轻核的实测丰度相一致.

从 1984 年以来,有若干研究组类似地讨论过理论的检验问题. 他们都乐观地认为原初核合成理论得到了证实. 但是对此的争议也一直没有间断过. 争论的实质在于所采用的原初丰度的实测推断值是否可信. 这责难无疑是有理的.

(6.27)中的下限来自 y_{23p}的上限. 如已指出,这上限的可靠程度并不是很清楚的. 如果它可靠,但是 y_{23p}的实际值比这上限小很多,那么相应的 η 值就可能太大,而与 Y_{4p}的实测推断相冲突(对照图 6.3 上的曲线看). 从另一方面看,(6.27)的上限来自实测推断的 $Y_{4p}<0.24$. 其他研究表明,Y_{4p}超过 0.24 完全可能,甚至超过 0.25 也并没有排除. 这样看来,现在认真检验核合成理论的条件尚不成熟. 实测确实显出对理论的支持,但是强烈地作肯定的断言则是过于乐观的.

值得注意类星体吸收云中氘丰度的测定. 潜在地,它可能是确定原初氘丰度的有力手段. 目前测到的低值($y_{2p}\approx 2\times 10^{-5}$)与 $Y_{4p}<0.24$ 是冲突的,而高值($y_{2p}\approx 2\times 10^{-4}$)则与 $y_{23p}<1\times 10^{-4}$相冲突. 这疑点的澄清,将是核合成理论检验上的重要进展.

从理论方面反思,上述理论中惟一可置疑之点是假定了核合成前质子数与中子数之比为均匀,而此前的夸克-强子相变(发生在 $T=100$ MeV 时)造成它不均匀是可能的. 然而已有研究表明,这种非均匀的核合成理论的结果与事实明显地难以符合. 这样,均匀的核合成几乎是惟一的可能. 因此值得强调: 原初的核合成过

程必定是发生过的，对这过程的理论处理又有十分可靠的基础，所以对其结果作全面检验未必是最重要的事. 在这理论得到充分检验前，我们也有理由相信它是对的，并相信从中获取的信息是可靠的. 它对中微子种数作出过正确的预言，这是一明证. 它对重子物质密度的预言也正受到越来越大的重视.

七、重子物质的密度

上面已讲到，在检验这些元素产额的同时，我们将能定出 η 的值. 它是研究原初核合成问题的一个极重要的副产品.

让我们简化地把 η 看成不随宇宙膨胀而变的量（实际上从那时到现在变化不大）. 那么就能从早期宇宙中定出的 η，来推出今天宇宙中由原子构成的普通物质的总密度. 这在宇宙学中是一个重要参量，被称作重子物质密度，记作 Ω_B. 无量纲密度是以临界密度 $\rho_c\left(=\dfrac{3H_0^2}{8\pi G}\right)$ 为单位衡量的. 记得 $\eta=n_N/n_\gamma$，于是可导出 Ω_B 和 η 的关系：

$$\Omega_B = \rho_N/\rho_c = 0.0037\eta_{10}h^{-2}, \tag{6.28}$$

其中 $\eta_{10}=\eta\times10^{10}$，$h$ 是以 100 km · s^{-1} · Mpc^{-1}为单位的无量纲哈勃常数.

如果相信原初核合成理论是不可能错的，那么得到任何一个原初丰度的实测推断，都能引伸出对重子密度 Ω_B 的预言.

从前面对原初丰度的讨论看到，由氘丰度的单调减小，得出 $y_{2p}>1.8\times10^{-5}$是最可靠的结果. 它预言的 $\eta_{10}<6.8$，按(6.28)有 $\Omega_Bh^2<0.025$. 再由 $h>0.5$ 推知

$$\Omega_B < 0.1. \tag{6.29}$$

这意味着若重子物质是今天宇宙的主要组分，那么宇宙肯定是无限的. 在得到这推论时，宇宙学家还没有以非重子为主的概念.

若接受 Walker 等人在 1991 年的结果，即上面讲到的 $\eta_{10}=$

2.8—4，再用 $h=0.5-0.8$，推得

$$\Omega_B = 0.016—0.059. \tag{6.30}$$

这结果对暗物质研究产生了很重要的影响.宇宙密度的局域测量已得到 $\Omega_0=0.2\pm0.1$，它明显地超过核合成理论预言的 Ω_B 值，构成了宇宙以非重子为主的强烈证据.这是以后要讨论的问题.

迄今为止，通过核合成研究定出 η，再由它推断 Ω_B 是宇宙学中确定重子物质密度的惟一方法.因为这结果对暗物质和结构形成研究都有重要影响，所以值得注意，目前的原初丰度推断还不能对重子密度给出很确切的预言.现在只能讲，Ω_B 的大小在 0.05 左右或更小一些.

第七章　物质与反物质的不对称

今天的宇宙中大量地存在着质子、中子和电子构成的物质，却没有发现有反质子、反中子和正电子构成的反物质．这是原始宇宙就有的不对称，还是宇宙演化而产生的现象？宇宙学有责任解答这个疑问．现在，解答疑问的线索已经找到，但是定量地成功的理论还没有出现．

一、有没有反物质天体

地球在宇宙中是一个很微小的局部，但这里反映出来的问题可不一定是局部问题．地球上的一切原子（或分子）都由质子、中子和电子构成，而没有相应的反原子（或反分子）．这种粒子与反粒子的不对称，就很可能是宇宙的全局特征，而不是局域现象．当然，作出这判断需要证据．

首先的问题是怎么区别其他天体是由正物质组成，或是由反物质组成．从微观讲，粒子和反粒子的电磁作用和强作用性质都一样，因此反质子和反中子构成的反原子核的结构与正原子核完全一样；再配合正电子而构成反原子，其结构与正物质原子也没有区别．若我们限于观测天体的电磁辐射，因光子和反光子又是同一种粒子，正物质天体与反物质天体是看不出差别的．要是我们能容易地接收中微子信号就好了．在同样的情况下，若正物质天体辐射中微子，则反物质天体将辐射反中微子．这个差别地面接收器能够区分．可惜今天我们还难以接收远处天体的中微子辐射．

能让我们判断地球之外是否有反物质天体的重要根据是粒子与反粒子相碰时的湮灭现象.湮灭过程将大量地释放出 π 介子.π^0 介子很快会衰变成光子,且这光子谱有很鲜明的特征.它是较易于发现的.

我们怎么知道太阳系内没有反物质天体?首先,太阳本身不是反物质星体,我们有直接的证据,因为太阳燃烧释放的中微子是已经测到了的.在太阳系的行星际空间充满着从太阳表面逸出的高速原子,这就是太阳风.它主要由质子组成.如果太阳系中有反物质行星或卫星,那么它不断受到太阳风的撞击,就会有持久而强烈的湮灭现象.没有观测到这种现象,表明整个太阳系内是没有反物质天体的.

同样的论据可类似地用于整个银河系.银河系的恒星际空间里充满着稀薄的星际气体.此外还有宇宙射线粒子在飞行.它们不与任何一个恒星相湮灭,构成了银河系内也没有反物质星体的有力证据.这种研究已延伸到本星系群的范围.因此天文学家已能肯定,在我们周围 10 Mpc 的范围内没有反物质星系.

更远处是否会有反物质星系?这问题尚不能由观测做确切的回答. 1998 年发射上天的 α 磁谱仪(Alpha Magnetic Spectrometer)原来在研究计划中叫它反物质谱仪(Anti Matter Spectrometer),简称都是 AMS.它的重要目的之一就是寻找更远处的反物质天体.

如果更远处有反物质星系,从那里逸出的高能反原子核将能自由地飞行到这里,被在地球大气层外运行的 AMS 接收到.注意接收到反氢核(反质子)是不足为凭的,因为它多半是宇宙线粒子碰撞产生的次级粒子①.如能接收到反氦核,意义就大不相同了.反氦核包含有两个反质子和两个反中子,它无法通过碰撞次级产

① 地面上接收到的宇宙射线主要来自银河系.宇宙射线中的反质子被认为是超高能射线粒子在中途碰撞产生的.

生. 因此 AMS 一旦接收到反氦核,那怕只有一例,就将是远处有反物质星系存在的有力证据. AMS 已上天工作了两年多,至今还没有令人激动的消息.

在理论家看来,10 Mpc 范围内没有反物质星系,几乎等于肯定了全宇宙没有反物质星系. 如果宇宙中正反物质是等量的,按粒子物理已经知道的相互作用,它们是没有可能作大尺度的分离的. 因此如果 AMS 证实远处有反物质星系存在,那对于物理学将是极大的挑战.

二、正反物质不对称疑难

综上所述,我们有理由相信,今天宇宙中没有大块的反物质. 正反物质是严重地不等量的,即所谓不对称的. 那么就引起了它怎么造成的问题.

往前追溯到原初核合成之前,今天没有反物质大体意味着当时的次要组分中只有少量的质子和中子,而没有同等的反质子和反中子. 我们关心的问题转化为:为什么当时会没有反核子?为回答这问题,还需要再往前追溯.

在宇宙温度远高于 1 GeV 时,正反核子(应表现为正反夸克)能热碰撞产生,因此当时反核子肯定也是大量存在的. 要紧的是正反核子等量或不等量. 待温度降至核合成要开始的温度,绝大部分正反核子已成对湮灭了. 湮灭后只剩下核子而不剩反核子,暗示着它们在湮灭前是不等量的. 让我们估算一下湮灭前的不对称程度.

在 $T\gg 1$ GeV 的超高温气体中,核子与反核子都是近似地没有静质量的辐射粒子. 它们的粒子数与光子数应同量级. 仅因强子的自旋自由度多一些,粒子数也多一些. 到了核合成时期,剩下的核子数只有光子数的 10^{-10}到 10^{-9}倍. 由此推断,湮灭前正反核子的不对称度极低. 用上一章得到的 $\eta=(3\sim 4)\times 10^{-10}$算,事先的不对称度 A 仅需为

$$A = \frac{n_b - n_{\bar{b}}}{n_b} \approx 3 \times 10^{-8}. \qquad (7.1)$$

在 1 亿个核子所占据的区域里，反核子的数目只比它少 3 个，就足以在湮灭后留下原初核合成所需要的核子数了. 推理出这结果很有启发.

在理论家看来，原始宇宙中正反核子没有理由不对称. 今天的不对称必定是宇宙演化的后果. 这样想来，在宇宙为超高温的甚早期应当有过一种物理机制，它使正反核子数目从完全相等变成了有微小的一点差别. 实现了这一步后，今天没有反物质天体就是自然的后果了.

那么正反核子数可能从相等变成不等吗？粒子实验的经验告诉我们：这不可能. 关键点是重子数有它的守恒性.

重子是核子概念的延伸. 它是质子、中子和其他更重的强作用费米子的统称. 人们按粒子实验的经验，引入了重子数 B 的概念：

任一重子(质子、中子等)具有	$B=1$,
任一反重子(反质子、反中子等)具有	$B=-1$,
一切其他粒子具有	$B=0$.

这里其他粒子包括了轻子、介子和规范粒子等. 一个粒子系统的总重子数 B 是组分粒子的 B 数的代数和. 实际上也就是重子与反重子的个数之差. 由实验总结出来的经验规律是：任何过程的前后，系统的总重子数 B 保持不变. 这就是刚才提到的重子数守恒律. 一对正反重子具有重子数 $B=0$. 它们必定成对地产生并成对地湮灭，正是 B 数有守恒性的体现.

用于宇宙，我们不能讲它的总重子数(可能是无穷大)，而只能讨论它的重子数密度，记作 n_B. 它是单位体积内的正反重子个数之差[①]，即

---

① 小心字面上的混淆. 对含有反重子的系统，重子数密度不是单位体积内的重子个数. 仅对不含有反重子的系统才是.

$$n_B = n_b - n_{\bar{b}}. \tag{7.2}$$

其中 n_b 和 $n_{\bar{b}}$分别代表重子和反重子的粒子数密度. 前面提到的正反重子不对称程度,可由 $A = n_B/n_b$ 描述.

考虑到重子数有守恒性,要在甚早期宇宙中发生从 $A=0$ 到 $A\neq0$ 的转变是不可能的. 不管所需要的不对称度有多小. 这样,理论家心目中的演化图景竟然是实现不了的! 宇宙甚早期中正反重子的微小不对称只能归之于宇宙的初条件,即在宇宙创生时正反重子的数目就稍微地差了一点. 这是理论家难以接受的结论. 所以把这问题称为正反重子不对称疑难.

三、重子数守恒是严格的规律吗?

重子数的守恒律是从大量粒子实验里总结出来的经验规律. 至今被研究过的粒子碰撞过程多得无法计数,但是没有发现过一个破坏重子数守恒的事例. 这的确把我们的问题引入了困境. 要是重子数守恒是一个严格的物理规律,那么我们遇到的宇宙学疑难是没有出路的. 人们只好被迫接受: 宇宙中正反物质"先天地"就是不对称的. 然而理论家从物理概念发展史的启示下意识到: 经验中没有破坏重子数守恒律的先例,并不一定表明它是严格的物理规律. 它依然可能只是在一定条件下才成立的唯象规律.

历史上,人们曾把化学性质不变的过程叫物理过程. 实际上指的仅是力学的或热学的过程. 在这种过程中,分子是物质的基本砖块. 它是守恒的. 深入到化学层次,结论就不同了. 化学过程是原子的重新组合,分子不再守恒. 原子作为化学过程的基本砖块,它表现出了守恒性. 值得强调,在无数次的化学实践中,绝不会有一个破坏原子个数守恒的事例. 炼金术家想把汞变成金,最后以失败告终,更强化了原子数守恒的观念. 那么,原子的守恒性是严格的规律吗?

在物理学深入到原子核的层次后,人们开始认识到,原子数的

守恒只是能量很低的过程中的唯象规律，而不是普遍规律. 如果过程中的能量达到核能量的范围，原子中的核与核能够直接相碰，那么化学元素是可以转化的，即原子(核)是不再守恒的. 这是一个很好的历史借鉴.

在认识到原子(核)可以不守恒的同时，人们发现了更基本的物质砖块，即质子和中子. 核反应是质子和中子的重新组合，相应的质子数和中子数是分别守恒的. 但是在核的β衰变中，质子与中子能相互转化，而并不分别守恒. 于是人们把质子和中子归为同一类粒子——叫核子. 由核过程的经验总结出：系统的总核子数(而不是总质子数或总中子数)保持守恒. 发现基元粒子须归类别才有守恒性，这是认识物质守恒规律的又一次深化. 重子数守恒的概念正是在这样的深化中产生的.

50 年代前半叶，随着加速器能量的提高，开始发现了超子 比核子更重而属同类的粒子. 例如出现了这样的过程，

$$\pi + p \to \Lambda + K^0, \tag{7.3}$$

过程前有一个核子，而反应后却不再有核子出现. 核子数守恒被破坏了. 当类似的过程越来越多，人们发现不是要废弃核子数的守恒律，而是需要把它拓宽. Λ 等一批新粒子与核子的共同点为：大家都是参与强作用的费米子(K^0 则是参与强作用的玻色子). 只是它们比核子重，所以被称为超子. 如果扩充核子的概念，把超子也包括进去，这新的大类(被称为重子)将依然具有守恒性. 人们就这样建立了重子的概念，并在至今的实验中，没有发现过违反重子数守恒的事例.

值得补充的是，如果重子数有严格的守恒性，那么最轻的重子只能是稳定的，因为不稳定粒子须衰变成比自己更轻的粒子. 若重子数可以不守恒，质子就可以衰变成(例如)正电子、再放出其他粒子. 质子的衰变没有被发现过，这就像汞变不成金一样，加深了人们对重子数守恒的信心.

现在让我们再考虑，有这么些实验证据的重子数守恒律一定

是一般的物理规律吗?上述分析说明,物质基本砖块的守恒方式是随着能量的提高而有变化的.因此我们完全不必过早地断言,重子数及其守恒在更高的能量下不会被更新的概念和规律所代替.现今的粒子实验只达到 10^3 GeV 的能量.在这能量以下,重子数的守恒是毋庸置疑的.但是当能量再提高很多量级,发现重子数的守恒律需要进一步拓宽是完全可能的.实际上,这样的试探性理论在70年代中期就开始出现了.

70年代初,弱作用和电磁作用相统一的理论得到了实验的证实.在它的启发下,有人开始想到,强作用与弱电作用也应是统一的.这就是大统一思想的发端.若强作用与弱作用是统一的,那么重子与轻子就没有了本质的差别,它们应可归入更大的一类.在它们之间发生转化(例如质子衰变成正电子)将是可能的.更广的守恒律将代替重子数(及轻子数)守恒律,对过程起着制约作用.要是这大统一的思想是正确的,重子数的守恒就不是普遍地成立的物理规律,而只是能量不太高的过程中的唯象规律.这样,宇宙学中的正反物质不对称问题也就不再是无法解决的"疑难"了.宇宙学家从这里看到了问题的出路,看到了曙光.

四、粒子与反粒子对称吗?

让我们回到理论家原来的思路上来.他们猜想,在刚创生的宇宙中,正反物质是等量的,后来的不等量是在宇宙演化中产生的.若这猜想是对的,上面已说明,自然界必定有能破坏重子数守恒的相互作用.大统一概念暗示了这类作用很可能是存在的.但是仔细分析发现,这只是一个必要条件,而并不是充分条件.进一步需要弄清,正反粒子的物理性质完全对称吗?若回答为"是",那么宏观的正反重子不对称也仍是演化不出来的.

为理解这道理,使用一些粒子物理术语是方便的.例如看最简单的过程:粒子 X 衰变成 Y 和 Z,即

$$X \to Y + Z. \tag{7.4}$$

假设它是一个使重子数增加的过程.某过程的粒子共轭(记做 C 变换),指把粒子与反粒子互换.按上例是(看图 7.1 上部)

$$\bar{X} \to \bar{Y} + \bar{Z}. \tag{7.5}$$

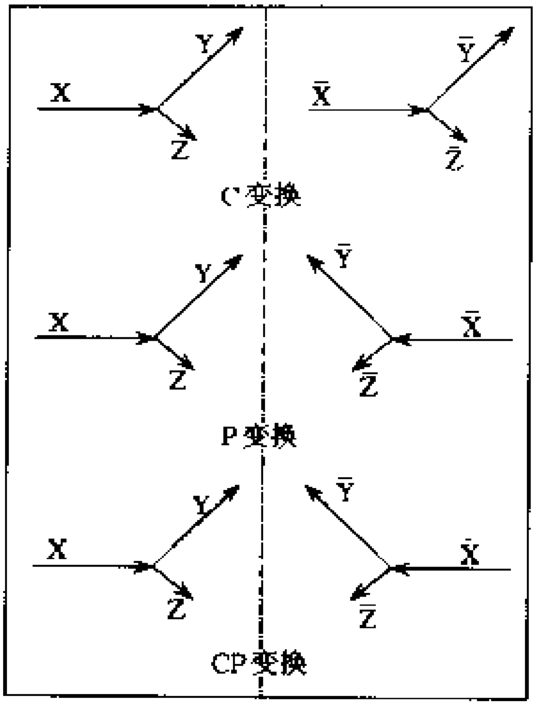


图 7.1　C 与 CP 变换

由于正反粒子的重子数必反号,所以后一过程必是使重子数减少的.现在要紧的问题在于这样两个过程的发生几率是否相同.用粒子物理的语言讲:如果发生几率相同,叫 C 守恒;若几率不同则叫 C 不守恒(或 C 破坏).让我们讨论几率相同与否的意义.

按照定义,粒子和反粒子必有相反的电荷,相反的重子数.其他能反号的性质也都相反.如果 C 守恒,即两者的相应过程都有相同的几率,那么正粒子与反粒子就完全是相对的.正与反的区分是没有客观标准的.反之,若 C 不守恒,我们就有客观标准来区分正粒子和反粒子了.因此 C 守恒或不守恒是告诉我们:粒子与反粒子的物理性质完全对称,或不完全对称.

回到正反粒子等量的原始宇宙上来.若一切粒子过程都满足

C 守恒，那么在超高温的宇宙气体中既发生着增加重子数的过程，也同时发生着减少重子数的过程，而且两者发生的几率也一样. 这是从微观角度讲的. 宇宙学感兴趣的是宏观效果. 重子数增加和减少的几率一样，就把宏观效果抵消了. 最终重子数等于零的状态还是不会改变. 即使只希望出现宏观上很微小的不等量，它仍是演化不出来的. 因此，期望由演化产生正反粒子的不等量，则必须有 C 破坏，即要求正反粒子的性质不是完全对称的.

这讨论还需要一点补充. 图 7.1 中部画出另一种变换，叫 P 变换. 它指把某过程换成其镜子里的过程. 这实际上是把左和右作替换. 这样两个过程有相同的发生几率叫宇称守恒（意即左右对称），反之就是宇称不守恒（即左右不对称）. 图的下部画的是 CP 变换. 它指把某过程替换成反粒子的镜像过程. CP 守恒与否可类似地理解. 我们现在需要补充的是：若 CP 保持守恒，按完全一样的道理，原初气体中使重子数增加的微观过程的效果将被其 CP 共轭过程的效果相抵消. 这样依然不能通过演化，使宇宙从 $n_B=0$ 的状态变成 $n_B\neq0$ 的状态. 因此要在宇宙中演化出正反重子的不对称，不仅需要 C 破坏，还需要 CP 破坏.

简单归纳一下上面分析的结论. 要是把今天宇宙中没有反物质星系看成是演化的结果，那么宇宙学就向超高能粒子物理提出了两方面的要求：（一）必须有破坏重子数守恒的作用存在；（二）在它起作用时必须同时有 C 和 CP 的破坏.

五、大统一与重子数的产生

上面的思想在 1967 年就由 Sakharov 提出了. 当时不仅实验上没有重子数可能不守恒的证据，而且理论上也没有看到这样的可能. 于是这想法被搁置了起来. 到 70 年代的后半期，在弱电统一的启发下出现了相互作用大统一的试探性理论. 在这种理论中，上面提出的两个要素是自然具备的，或是能够容纳的. 这样，用演化

来产生重子数的定量研究就开始活跃起来了. 由于这种努力最后没有成功,因此我们不进入细节,而只讨论它所描绘出来的演化图景. 它对我们具体地理解上面的要求是有帮助的.

设在大统一的能量尺度(10^{15} GeV)上存在很重的 X 介子及其反粒子 $\bar{X}$,其衰变是破坏重子数守恒的. 当宇宙温度降至 $T<m_X$,X 和 $\bar{X}$ 粒子将通过衰变而消失. 让我们看它们全部衰变后的宏观效果.

令 X 粒子有两种衰变方式,它们是

$$\begin{aligned} &X \to q+q, \qquad \text{分支比}^{①}\text{为 } r, \\ &X \to \bar{q}+\bar{l}, \qquad \text{分支比为 } 1-r. \end{aligned} \tag{7.6}$$

衰变终态中的 q 和 $\bar{q}$ 代表夸克和反夸克,其重子数 B 各为$\pm 1/3$. $\bar{l}$ 代表反轻子,其重子数自然是 0. 这样,X 粒子按前一方式衰变后,重子数增加了 2/3;按后一方式衰变后,重子数减少了 1/3. 平均每个 X 介子衰变产生的重子数为

$$B_X = r(2/3) + (1-r)(-1/3). \tag{7.7}$$

$\bar{X}$ 粒子必能衰变成相应的反粒子,即

$$\begin{aligned} &\bar{X} \to \bar{q}+\bar{q}, \qquad \text{分支比为 } \bar{r}, \\ &\bar{X} \to q+l, \qquad \text{分支比为 } 1-\bar{r}. \end{aligned} \tag{7.8}$$

其中 l 是轻子,其重子数也是 0. 因此每个 $\bar{X}$ 粒子衰变平均产生的重子数为

$$B_{\bar{X}} = \bar{r}(-2/3) + (1-\bar{r})(1/3). \tag{7.9}$$

把(7.7)和(7.9)相加,得到每对 X 和 $\bar{X}$ 粒子衰变平均产生的重子数. 它是

$$\epsilon \equiv B_X + B_{\bar{X}} = r - \bar{r}. \tag{7.10}$$

然后我们讨论全部衰变后的宏观效果.

在温度 $T>m_X=m_{\bar{X}}$的热平衡气体中,X 和 $\bar{X}$ 的数密度相等. C 或 CP 的守恒表明 r 与 $\bar{r}$ 相等. 这样在一切 X 和 $\bar{X}$ 都衰变后,所

① 分支比指按这种方式衰变的百分比.

产生的净重子数将是零. 这就是前面讲的,破坏重子数守恒的微观过程将不产生宏观效果. C 和 CP 都不守恒则有 $r\neq\bar{r}$,即 $\epsilon\neq0$. 这样才会有重子数产生的宏观效果出现. 在这意义上,参量 ϵ 可作为相应过程中 C 和 CP 破坏程度的标志.

过程进行的图景是这样. 虽然 X 和 $\bar{X}$ 是不稳定粒子,但当宇宙温度 $T\gg m_X$,X 与 $\bar{X}$ 能衰变也能热碰撞产生. 在热平衡下,其数密度 n_X 和 $n_{\bar{X}}$ 仍由(4.15)给出. 它与当时光子或重子的数密度都是差不多相等的. 从 $T\approx m_X$ 时起,热碰撞产生 X 和 $\bar{X}$ 的几率减小,它们的数密度开始因其衰变而减少. 气体中的净重子数就开始产生. 理论上要计算的是,至 X 和 $\bar{X}$ 完全衰变,宏观地产生了多大的净重子数密度. 因为在这衰变阶段,X 和 X 的数密度并不始终满足热平衡分布,所以必须用非平衡统计的玻尔兹曼方程来计算其数密度的变化. 我们不想涉及复杂的计算细节,而只作很粗略的估算.

让我们把所有 X 和 $\bar{X}$ 粒子的衰变看成是瞬时完成的. 衰变前的数密度都是 n_X. 那么在它们完全衰变后,所产生的重子数密度 $n_b-n_{\bar{b}}$ 为 ϵn_X. 气体中的正反重子不对称程度 A 达到了

$$A=\frac{n_b-n_{\bar{b}}}{n_b}=\epsilon\frac{n_X}{n_b}\sim\epsilon. \qquad (7.11)$$

这样看来,所产生的不对称度主要是由 C 和 CP 的破坏程度决定的. 当然要注意,如此粗略的估算与实际结果可有量级的差别. 具体地研究某种模型时,必须用玻尔兹曼方程认真地处理它的衰变阶段.

在粒子物理学家试探地建造大统一理论时,格外地被关注的是以 $SU(5)$群为基础的模型. 人们猜想,最简单的可能正是对的. 最简单的 $SU(5)$大统一模型中包含的自由参量最少,其理论预言也粒确定. 这模型中自然地包含有破坏重子数守恒的相互作用,也有相应的 C 和 CP 破坏. 因此它为研究重子不对称起源提供了理想的出发点. 20 世纪 80 年代前后,人们按上面描述的图景,具体

地计算了模型中的规范玻色子或 Higgs 粒子(作为 X 粒子)衰变时产生的重子不对称. 遗憾的是在参量的合理选取下,所得到的结果与核合成的需要差了许多量级.

六、理论的现状

大统一理论的出现,很自然地使重子数的演化产生有了可能. 因此定量的结果不对说明什么? 这是需要理论家思考的问题.

按 $SU(5)$大统一模型,质子应也是不稳定粒子. 估出的寿命是 $\tau_p \sim 10^{30}$年. 能否用观测来证实它,当然很关键. 质子的寿命这么长,要发现它的衰变很难. 无论如何人们意识到,在一个空的矿井中对周围大量质子作搜寻,在不太长的时间里发现一个或几个质子衰变事例是可能的. 于是若干这样的实验装置建立了. 经过好多年的观测,结果没有发现一个质子衰变事例,从而推出质子寿命的下限是 10^{32}年. 它与 $SU(5)$大统一模型也不符. 这不符告诉宇宙学家:用它不能得到足够的重子数产生是由于这模型本身不对. 因此,宇宙中正反重子不等量的起因需要更深入研究.

我们知道,大统一的能量尺度太大,这能量范围的物理规律尚很不清楚. 人们到 80 年代末期发现,正反物质的不对称也会在弱电作用的能量尺度($\sim$200 GeV)上产生. 这是一个重要的进展,因为这样把宇宙学非常需要发生过的过程降到有可靠物理基础的范围内了.

从弱电统一理论本身看,它的拉格朗日量中不包含能改变重子数的因素. 重子数在这能量下的变化是通过隧道效应出现的. 弱电的统一作为非阿贝尔规范理论,其真空(叫 θ 真空)有复杂结构. 图 7.2 示意地画出了它. 图上的纵轴是自由能密度,横轴是规范场或 Higgs 场的场量. 不同的真空间有势垒相隔. 当粒子从一个真空穿入另一相邻的真空,重子数的改变为 $\Delta B=3$. 这样当温度较高,隧道效应将造成重子数的变化.

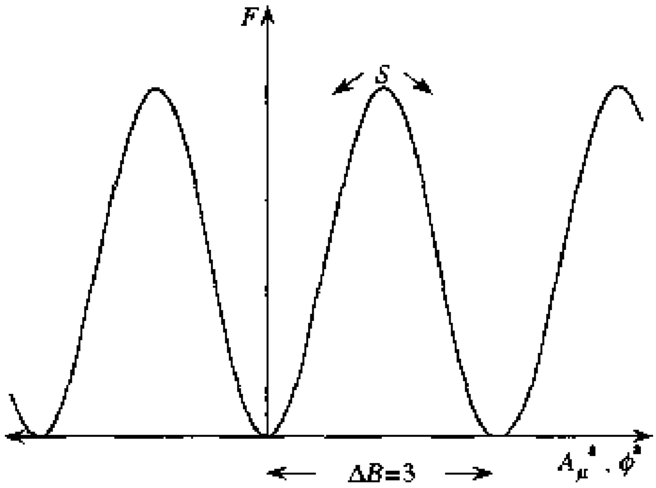


图 7.2　θ 真空结构的示意

今天的粒子物理把弱电统一,而强作用未统一在内的 $SU(3)\otimes SU(2)\otimes U(1)$模型叫标准模型. 其中 $SU(3)$描述强子的色相互作用,$SU(2)\otimes U(1)$描写统一的弱电作用. 对于能量低于 10^3 GeV 的现象,实验已证明它是很对的. 理论家原则上认识到,超高能现象会不服从这标准模型,它仅是"最终"理论的低能近似. 因此人们感兴趣于发现与标准模型有冲突的现象. 这是进一步建立超高能理论的需要.

在标准模型基础上按上述机理计算重子产生表明,CP 破坏程度不够. 因此为了得到实际的正反物质不对称度,必须找到新的 CP 破坏机制才行. 此外,Higgs 粒子尚未发现是标准模型的重大遗留问题. 实验定出的 Higgs 粒子质量下限为 100 GeV,而由重子数产生研究给出的质量限仅为 40 GeV. 因此把粒子物理标准模型用于宇宙,人们也看到这标准模型需要扩充,以包含更多的性质. 例如中微子应当有质量,玻色子与费米子间应有超对称等. 宇宙学的研究无疑需要依靠物理学作为基础. 这里看到,现在前者也正在为后者的发展提供有益的启示.

正因为宇宙的重子产生问题不能用已成熟的粒子理论解决[1]，可以料想，我们不可能在短期内弄清正反物质不对称的实际起因. 这宇宙学问题将与超高能理论研究同步地发展，并相互地提供借鉴.

---

[1] 张新民教授详细地为作者分析过在弱电尺度上做重子产生的收获与困难. 谨此致谢.

第八章　甚早期的暴胀

宇宙甚早期曾短暂地发生过剧烈的加速膨胀，这是宇宙学研究中最重要的概念之一. 这思想使原来无法解释的视界疑难和准平坦性疑难等得到了自然的解决. 更为重要的是，它预言了今天的宇宙密度 Ω_0 很接近于 1，因此非重子物质才是介质中最主要的组分. 此外它也为原初扰动的起源提供了线索. 这些后果对宇宙学的发展产生了深刻而积极的影响. 可惜的是对暴胀本身的实际起因，今天尚没有肯定的答案.

一、标量场的自作用

粒子物理中，相互作用的统一是借助对称性的自发破缺而实现的. 80 年代初，在这一思想的启示下，Guth 提出了宇宙甚早期发生过暴胀的理论. 因此在研究暴胀之前，须先讨论相互作用统一和对称性自发破缺的机理.

在微观粒子的相互作用中，电磁作用被研究得最早，也最清楚. 用量子规范场的语言讲，电磁场是一种规范场. 它与由 $U(1)$ 群描述的内部对称性相联系. 弱作用在表观性质上与电磁作用很不同. 电磁作用是长程力，而弱作用的力程极短. 就强度而言，两者差很多量级. 因此人们原来把它们看作两种本质上不同的作用力.

弱电统一理论发现，这两种力可统一地用具有 $SU(2)\otimes U(1)$ 对称的规范场来描写，即它们原是同一种力. 理论的要点是，这对

称性在低能条件下会自动地出现破坏，从而使弱力和电磁力表现得很不一样．只有当能量很高，原有的对称性才恢复，两种力的性质才显出完全一样．70年代以来的大量粒子实验已高度证实了这理论的许多预言．

在弱电统一理论中，对称性的自发破缺是靠一种带自作用的标量场 ϕ 来实现的．为定性地说明这机理，设其自作用势有形式

$$V(\phi) = -\mu\phi^2 + \lambda\phi^4, \tag{8.1}$$

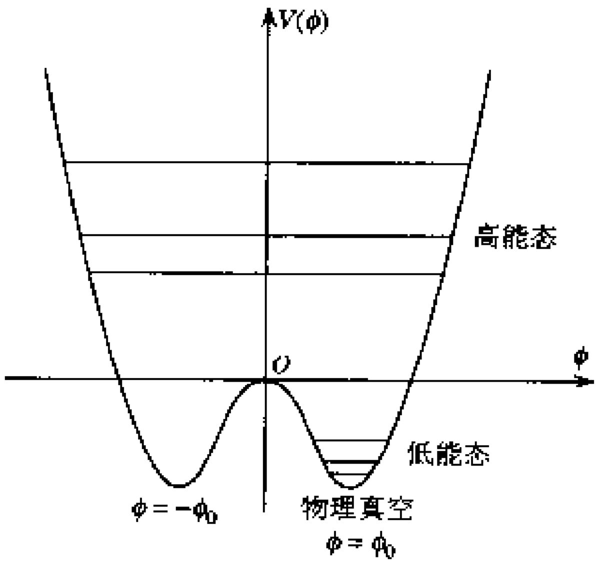


图 8.1　Higgs 位势

其中 μ 和 λ 是实参量．这简化的例子中，ϕ 场只具有一种最简单的内部对称，即 ϕ 与 $-\phi$ 互换不变．图 8.1 是这个位势的示意图．横轴代表场量 ϕ，纵轴是相应的势能密度．这势能对 $\phi=0$ 有左右对称．

场论中把能量密度最低的状态叫基态，或真空态．对于有上述自作用的 ϕ 场，$\phi=0$ 处是势能的极大而不是极小，因此不代表真空．从图上看出，真空态是 $\phi=+\phi_0$ 或 $-\phi_0$．当研究低能现象，即过程的能量在 $V(\pm\phi_0)$附近，由于两种真空之间有势垒相隔，不能同时实现．现实的真空只能是两者之一．图上设它为 $+\phi_0$．这样在系统的物理状态中，ϕ 与 $-\phi$ 的互换对称性自动消失了．这就是对称性自发破缺的 Higgs 机制．相应的 ϕ 场叫做 Higgs 场．从图上同样

看到，当所研究系统的能量很高，ϕ 与 $-\phi$ 的互换对称性会自动恢复. 它正是弱电统一理论所要的效果.

弱电统一理论中的 $SU(2)\otimes U(1)$ 对称的含义比上述简单例子要丰富，但是难以用图画出. 那里对称性的自发破缺是部分地使对称性得不到体现，而不是对称性的全部消失. 我们不需要进入细节. 高能和低能的界限是由 $\phi=0$ 处和 $\phi=\pm\phi_0$ 处的势能之差标志的. 对于弱电统一理论，这分界能量是 10^2 GeV. 按当时的实验条件，这能量超过了加速器所达到的范围. 现代大加速器的能量已达到 10^3 GeV.

在各种相互作用统一理论中，内部对称性的自发破缺是一个基本概念. 大统一理论也使用这个概念. 它把强作用和弱电作用统一地看成与更大内部对称群(例如 $SU(5)$)相联系的规范场. 然后要求这个对称群随能量降低而逐级地自发破缺. 当能量高过 10^{15} GeV，强、弱、电作用才是实际上统一的. 在能量低于这能限，物理上将看到强和弱电两种力. 当能量再低过了弱电统一的能限(10^2 GeV)，我们才看到强、电和弱三种性质不同的力. 把自然界一切力都归结为一种统一的力，无疑是很美好的思想. 但是这理论至今尚没有得到实验的证实.

我们感兴趣的是这思想对宇宙学的影响. 从宇宙学看，今天很冷①($T=3$ K)的宇宙必是从早期很热($T>10^{15}$ GeV)的状态下演化过来的. 因此能量很高时应发生过的对称性破缺肯定对甚早期高温宇宙产生过影响. 本章要讨论的暴胀正是这种影响之一.

二、真空相变

研究早期宇宙时我们面对的是高温介质. 因其组分是高能粒子，所以高能物理规律自然是必要的研究基础. 但是粒子物理与宇

① 今天宇宙中的实物已没有统一的温度，只有光子背景才仍有统一的温度.

宙学有所不同.粒子物理处理的是少数粒子的动力学系统,而宇宙学处理的是大量粒子构成的统计系统(即气体).为考虑对称性自发破缺的宇宙学效果,我们须讨论高温均匀介质中的 Higgs 场.

Higgs 场 ϕ 在温度为 T 时的平衡态,由其自由能密度 $F(\phi,T)$ 决定.零温的统计系统就是纯动力学系统,因此零温 ϕ 场的自由能密度 $F(\phi,0)$ 就是它的自作用势函数 $V(\phi)$.任意温度下的 $F(\phi,T)$ 可由 $V(\phi)$ 出发,用量子统计方法算出.类似于粒子的平衡位置必使其势能 $V(\phi)$ 取极小值,温度为 T 的 ϕ 场平衡态必使相应温度下的自由能密度 $F(\phi,T)$ 取极小.这状态就叫温度为 T 的真空态.让我们讨论 ϕ 场的真空态随温度 T 的变化.

当有如(8.1)式所示的自作用势,算出的 $F(\phi,T)$ 随 T 的变化定性地画在图 8.2 中.问题中有一特征温度 T_c.当介质温度 T 高于 T_c(见图 8.2(a)和(b)),Higgs 场的真空态为 $\phi=0$.这真空态已恢复了系统应有的(ϕ 与 $-\phi$ 互换)对称性,因而被称为对称真空.在温度降到接近 T_c 时,自由能 F 开始在 $\phi\neq0$ 处出现新的极

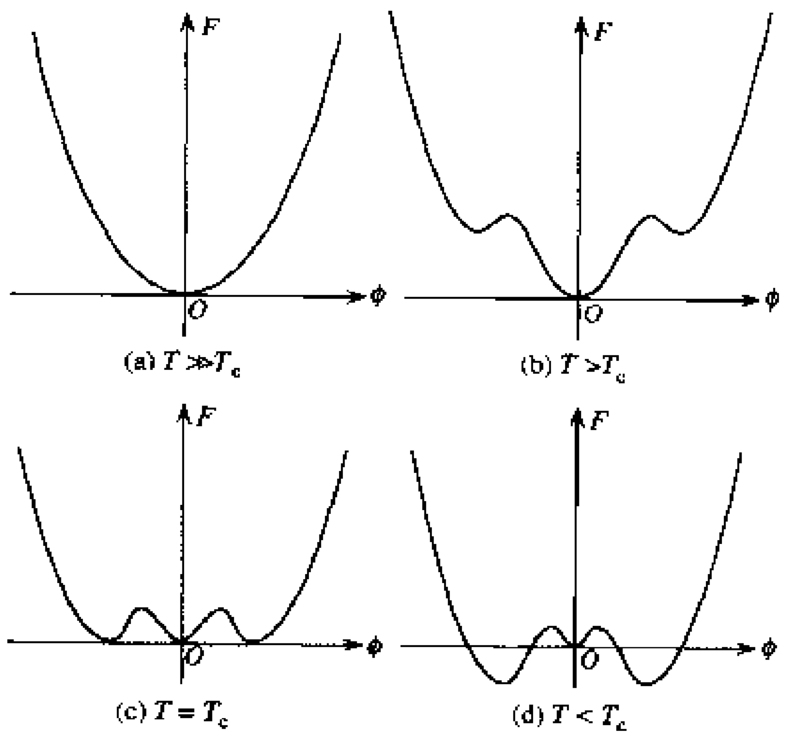


图 8.2　自由能随温度的变化

小.相对于 $\phi=0$ 的真空,它们的自由能较高,因此是亚稳的假真空.图 8.2(c)画出 $T=T_c$ 的临界情形.这时真假真空具有相同的自由能,即若干真空态简并了.它是真空将发生相变前的临界状态.当温度降到低于 T_c 时,情况如图 8.2(d)所示.这时 $\phi\neq0$ 处的极小是物理真空,而 $\phi=0$ 处的极小成了亚稳的假真空.$\phi\neq0$ 的真空已不再具有 ϕ 场原应具有的(全部)对称性,因而被称为对称破缺真空.这些结果使我们看到,当介质从 $T>T_c$ 降温到 $T<T_c$,真空态要发生一次突变.这就是对称性自发破缺引起的真空相变.

真空相变的临界温度 T_c 当然取决于所处理的问题.对破坏弱电统一的真空相变,T_c 约是 10^2 GeV. 对破坏大统一的真空相变,T_c 将是 10^{15} GeV 的量级.我们把 $T=0$ 的(对称破缺)真空规定为能量的零点,则临界真空态的能量密度 $\rho_{vac}\approx T_c^4$.

现在我们把这真空相变的概念应用于随膨胀而降温的宇宙.

当宇宙温度降到 T_c 后,这时应发生真空态从 $\phi=0$ 向 $\phi=\phi_0$(或 $\phi=-\phi_0$)的跃迁.但是从图上可看出,这两个真空态之间隔着很宽的势垒,跃迁不能立即发生.但这不妨碍宇宙的继续膨胀,从而气体的温度在继续降低.用统计热力学的语言讲,真空进入了过冷状态.这时真空的能量密度仍保持为 $\rho_{vac}\approx T_c^4$,而这真空背景上的辐射气体的能量密度却按 $\rho_\gamma\approx T^4$ 的规律下降.因此当降至 $T\ll T_c$,宇宙总密度将主要来自真空的贡献.气体本身变得越来越次要,以至可以忽略不计了.用通俗的话讲,宇宙中原有的物质都被洗掉了.这样的局面将维持到真空相变实际发生的时刻.下节中将指出,在相变实际发生之前,宇宙将按指数律而剧烈膨胀.

当宇宙温度降到一定程度,从过冷的对称真空向对称破缺真空的相变总要发生的.由于在发生时 ϕ 场的两种真空已有很大的能量差,这能量差将作为相变潜热放出.ϕ 场与所有其他场都有相互作用.通过这些作用,ϕ 场的假真空能将转化为其他场的能量,

使其他场受到激发①.场的高度被激发也就是大量地产生场量子.在达到热平衡后,等于重新建立了一个高温气体.从能量守恒看,这气体的温度将大致与 T_c 相当.由于新建的气体温度远高于相变前的温度,这过程叫宇宙的重加热.完成这重加热后,新的真空能又远小于气体的能量了.这时真空能可忽略,宇宙又恢复了以辐射为主的"正常"状态.

三、暴胀的产生

早期宇宙的膨胀规律已在第四章第四节中讨论过了.略去曲率项的影响,动力学方程的形式是

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G}{3}\rho, \tag{8.2}$$

其中的 ρ 是宇宙的总密度.原来人们把物质分成辐射和实物两部分.早期宇宙是以辐射为主的,实物的贡献被忽略.现在看到有的时候真空能密度的贡献会很重要,因此在研究早期动力学时,方程中的 ρ 应把它包括进去,即

$$\rho = \rho_{vac} + \rho_\gamma. \tag{8.3}$$

先讨论 $T \gg T_c$,即相变之前.这时虽然 $\phi = 0$ 的对称真空有很大的能量密度 T_c^4,但是只要 $T \gg T_c$,宇宙中高温辐射气体对密度的贡献占压倒优势,真空能可忽略.所以原来的标准模型依然成立.我们已经知道,宇宙在这情况下将按

$$R = At^{-1/2} \tag{8.4}$$

的规律正常地膨胀.

重要的情况发生在相变前.由于相变并不能在 $T = T_c$ 时及时地发生,宇宙进入了亚稳真空的过冷态.宇宙的继续膨胀使辐射气体的密度($\sim T^4$)迅速下降,而真空能密度 T_c^4 却不变.于是出现了

① 例如电子场处于真空表现为没有电子.电子场受到激发,就是电子的出现.

宇宙总密度中以真空为主的新情况.这时动力学方程为

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G}{3}\rho_{\text{vac}}. \tag{8.5}$$

考虑到 $\rho_{\text{vac}}(\approx T_c^4)$为常数,动力学方程的解有形式

$$R \propto e^{Ht}. \tag{8.6}$$

其中

$$H = \left(\frac{8\pi G}{3}T_c^4\right)^{1/2}, \tag{8.7}$$

它表明在真空为主阶段,宇宙是按指数律膨胀的.这是标准模型所没有考虑到的情形.

让我们以大统一破缺相变为例来看这阶段膨胀的剧烈程度.代入 $T_c = 10^{15}$ GeV,可估出指数因子中 H^{-1}的大小.用普通单位制表示,它是

$$H^{-1} = 10^{-35}\ \text{s}. \tag{8.8}$$

由此看来,如果真空为主阶段持续了 10^{-33} s,则在这短暂阶段中宇宙尺度因子 R 增大了 e^{100}(即 10^{43})倍.这比标准模型在 150 亿年内膨胀的倍数①还多.所以人们把这种惊人的膨胀叫暴胀.

在真空态完成了从 $\phi = 0$ 到 $\phi = \phi_0$(或 $\phi = -\phi_0$)的跃迁,新的(对称破缺)真空的能量密度变得相对很低.紧接着的重加热过程把气体的温度又升得很高.这样,新产生的辐射气体又成了宇宙密度的主要来源.于是,宇宙的膨胀又恢复到了辐射为主的正常状态.相变前期的暴胀仅是一个短暂的插曲.

为示出暴胀对演化的影响,图 8.3 中用对数刻度示意地画出尺度因子 R 随时间的变化过程.今天的 R_0 被规定为 1.实线画的是有暴胀影响的模型.虚线代表没有暴胀的标准模型.在暴胀前或后的正常膨胀阶段,$\log R$ 与 $\log t$ 成正比;而在暴胀中,$\log R$ 与 t 成正比.图上假定暴胀从 $t = 10^{-35}$ s 开始,持续了 10^{-33} s,则 $\log R$ 增

---

① 这倍数是今天与普朗克时刻相比而言的.

加[1]了 43. 我们从图上看出，考虑到暴胀的影响，宇宙在暴胀前的尺度比按标准模型所想象的要小 43 个量级. 设把普朗克时刻($t=10^{-43}$ s)作为宇宙膨胀的起点. 如果没有暴胀，原初宇宙的尺度因子为 10^{-31}. 若有过这样的暴胀，原初宇宙的尺度因子应为 10^{-74}. 两个结果有巨大的量级差别. 这差别将使许多疑难问题得到自然的解决.

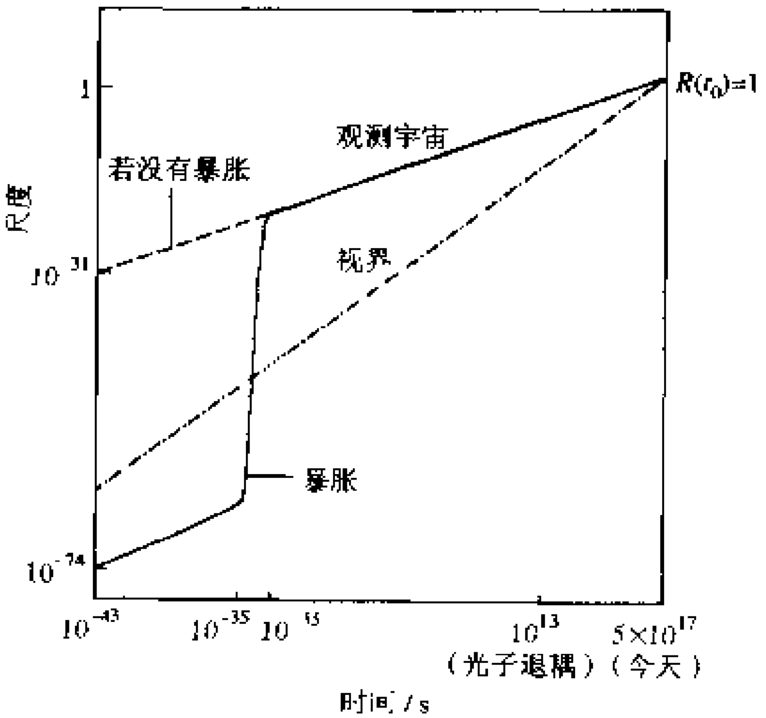


图 8.3 暴胀对尺度变化的影响

图 8.4 画的是宇宙温度 T 随时间的变化. 在相变前的暴胀中，因真空进入过冷态而使气体的温度陡然下降. 当真空相变完成时，相变潜热的放出使气体得到了重加热. 其温度又重新回到 T_c 附近. 此后真空能再次变得不重要. 宇宙温度将继续按标准模型的行为下降. 因此暴胀只影响了短暂阶段的温度变化.

---

[1] 这里用的是以 10 为底的对数.

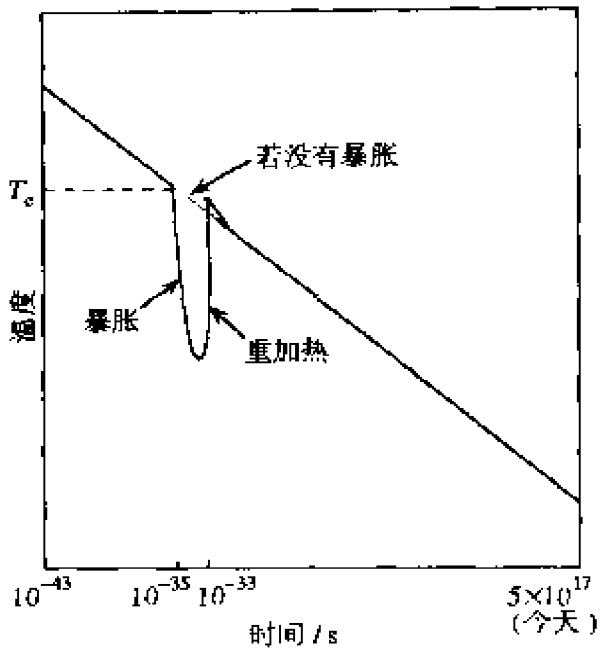


图 8.4　暴胀对温度的影响

四、均匀性疑难

在第四章中讨论宇宙视界的概念时曾提到过，视界的有限，特别是早期视界很小会引起全宇宙的均匀性无法解释. 但是全宇宙的均匀性终究只是理想化的理论假设. 当没有证据表明宇宙中密度均匀的区域显著大于视界，这疑问显得有点学究气. 然而，背景辐射温度的各向同性已现实地把上述疑问变成了尖锐的困境.

我们知道，观测到的背景辐射来自最后散射面(参看图 5.2)，从那里“发出”这光的时间是 $t_1=10^5$ 年. 那时的视界肯定很小. 不难计算出那时的视界大小，并把这样大小的区域放在最后散射面上(参看图 5.1). 它对我们的张角只有 1°. 这就是说，整个最后散射面的面积比视界区的截面积大几万倍！于是问题出现了：最后散射面上包含几万个不可能有因果联系的区域，为什么各处的温度会均匀？这正是早期视界太小造成的均匀性疑难.

为理解这疑难的深刻性，可以分析一下视界尺度和观测宇宙

尺度[1]的关系. 图 8.3 所画的就是观测宇宙大小随时间的变化. 同一图中已画上了视界的变化(点划线). 因为光的传播肯定比宇宙介质的膨胀要快,所以视界比 $R(t)$ 的增长更快是意料之中的事. 这样就导致了一个后果: 我们的观测宇宙在以前任一时刻的大小都比当时的视界要大,且越早大得越多.

具体到背景辐射形成(即作最后一次散射)的时刻,观测宇宙已比当时视界在线度上大几百倍. 论面积或体积,前者比后者则分别大了几万或几百万倍. 这范围内温度的均匀性成了物理上不可思议的事. 这是"现实的"均匀性疑难. 若再在时间上前推,差别更强烈. 推到普朗克时间,即 $t=10^{-43}$ s,观测宇宙的线度比当时的视界约大了 40 个量级! 各部分物质完全没有联系的可能,它当时怎么会是均匀的呢? 这是"放大了的"均匀性疑难.

在考虑到宇宙曾在甚早期的某阶段有过暴胀,这疑难就不存在了.

上节中已指出,若有过暴胀,观测宇宙在此前的大小远比原来估计的要小 43 个量级. 可是在暴胀阶段,视界的大小却是不变[2]的. 因此视界大小的变化如图 8.3 上所画,它不受暴胀的影响. 这样再看暴胀前两者的相对大小,则此前的观测宇宙比当时的视界要小几个量级(见图 8.3). 这样就没有疑难了. 观测宇宙内密度和温度完全可能在暴胀前通过扩散或其他物理过程而均匀化. 暴胀后,观测宇宙越出了视界,变成了内部各部分无法联系的系统. 但是各部分均匀的事实将保留下来. 这就是背景辐射有各向同性的原因,也是今天宇宙中各部分密度很均匀的原因.

从这样的分析看,存在过暴胀能使宇宙均匀性问题不构成疑

---

① 把观测宇宙理解为今天视界内的那些物质. 观测宇宙尺度指那些物质所占据的范围大小. 那么它当然是与 $R(t)$ 成正比的.

② 前面只讨论了光子传播时间有限性引起的"粒子"视界. 暴胀阶段的视界是由"事件"视界决定的. 我们没有仔细讨论过. 按事件视界的概念算出,暴胀阶段它不增大.

难，但是暴胀阶段须持续一定的长度. 理论上暴胀会持续多久，取决于理论模型的细节. 今天并不清楚. 原则地讲，它必须使 $R(t)$ 增大 40 个量级以上，否则暴胀前的观测宇宙依然比视界小，即均匀性疑难将依然存在.

五、结构起源疑难

第一章中已作为基本事实指出，今天的宇宙物质是层次性地结团的. 结团的原因是早期均匀介质中的微小密度起伏演化成了局域的大起伏. 结构形成问题将在最后一章作专题讨论. 现在只讨论原初密度起伏的起源问题. 我们以星系为例来讨论.

中等星系的质量为 $10^{10}M_\odot$，它们的典型大小只有 10 kpc. 若把今天的宇宙物质完全均匀化，星系质量的物质所占据范围的线度约是 1 Mpc. 它比整个观测宇宙小 4 个量级. 追溯到结团之前，这比例当然是不变的. 我们已经知道那时的观测宇宙远大于视界，那么准备形成星系的那块物质也必会大于当时的视界. 后来有星系的形成，需要早期有星系尺度的正密度扰动，即在 1 Mpc 范围内各处的密度都超过平均密度. 这种扰动就是形成星系的“种子”. 可是当其中各部分不能有物理联系，就不可能存在一种机制来产生这样的密度起伏. 于是早期视界过小又导致了结构的起源疑难：没有物理的方法能产生星系形成的种子，那么后来的星系是从哪儿来的. 当然这问题对星系团或超团也同样存在.

既然疑难的本质还是在于早期的宇宙视界太小，暴胀理论肯定也能使这问题得到排除. 重要的是，暴胀还为问题的答案提供了启示.

虽然星系尺度的变化没有在图 8.3 中画出，我们仍可从这张

图上来看[1]. 在早期很长一段时间里，星系尺度是超视界的. 但是当追溯到暴胀之前，整个观测宇宙已比视界小，星系尺度自然也成亚视界的了. 这也就是说，星系和各种尺度的结团种子在暴胀之前是可能产生的.

这里还有一个需要注意之点. 前面已提到，暴胀把原来的气体都稀释掉了，因此此前气体中的密度起伏也一起被洗掉了. 它们成不了后来结构形成的种子. 这样分析起来，后来成长起来的种子只能产生在暴胀的过程之中. 现在这已是多数研究者的共识. 如何正确地产生结构形成所需的种子是暴胀研究中的重要课题.

六、准平坦性疑难

让我们再讨论一个在暴胀理论面前会自然消失的困难问题，即准平坦性疑难. 它问：为什么我们的宇宙在演化了 100 多亿年后仍如此地接近于平坦？这问题本身需要先解释一下.

宇宙空间是否平坦是由曲率参量 k 标志的. 宇宙密度 Ω 直接与 k 有关，因此它的值也能反映空间的弯曲程度.

把 Friedmann 方程(3.1)看成任意时刻的代数关系，我们有

$$1-\frac{1}{\Omega}=\frac{k}{8\pi G_{l}\rho R^{2}}. \tag{8.9}$$

这里的 ρ 是把真空能也包括在内的宇宙总密度. 若 $k=0$ 即宇宙为平坦，则 Ω 永远等于 1. 当 $k\neq0$，由上式看到，

$$\left|1-\frac{1}{\Omega}\right|\propto\frac{1}{\rho R^{2}}. \tag{8.10}$$

现在用 $1-1/\Omega$ 作为空间偏离平坦程度的标志，分析它随时间的变化.

先忘掉暴胀，则早期宇宙以辐射为主，有 $\rho R^4=$常数；后期以

[1] 星系尺度也与 $R(t)$ 成正比. 只不过它比观测宇宙小，所以整个变化曲线向下平移即是.

实物为主，有 $\rho R^3=$常数. 这样看来，

$$\left|1-\frac{1}{\Omega}\right| \propto \begin{cases} R^2 & \text{（辐射为主），} \\ R & \text{（实物为主）.} \end{cases} \tag{8.11}$$

在整个宇宙演化期间，空间对平坦的偏离是单调地增大的. 从普朗克时刻算起到今天，它已增大了 58 个量级. 这样就出现了开头提出的问题：$1/\Omega$ 对 1 的偏离已有了 58 个量级的放大，为什么今天的 Ω_0 仍为 1 的量级？

这问题也可以反过来提. 如果宇宙不严格地平坦，从今天 $\Omega_0\sim 1$ 的事实往前推，那么在普朗克时刻应有

$$\Omega = 1 \pm O(10^{-58}). \tag{8.12}$$

于是问题转化为：原初宇宙的密度 ρ 与膨胀速率 H 怎么会有如此巧妙的配合，以致 Ω 在第 58 位小数上才有对 1 的偏离. 这需要细微调节才能出现的初条件极不自然. 因此被称为准平坦性疑难.

甚早期暴胀的发生对这问题的影响也很大. 暴胀阶段宇宙以真空为主. 真空能密度是个常数，使得在这短暂阶段有

$$\left|1-\frac{1}{\Omega}\right| \propto R^{-2}. \tag{8.13}$$

$|1-1/\Omega|$ 要急剧地减小. 如果像解决平坦性疑难所需，尺度因子 R 在暴胀阶段增大了 43 个量级，那么 $|1-1/\Omega|$ 在这阶段降低了 86 个量级. 从普朗克时刻算起，抵消掉暴胀前后上升的 58 个量级，$|1-1/\Omega|$ 还降低了 28 个量级. 原来的准平坦性疑难就不复存在了.

在这问题上，暴胀理论不仅解决了疑难，而且提出了新的预言. 我们合理地设想在普朗克时刻的 $1/\Omega$ 与 1 有任何“普通大小的”偏离，在降低了很多(例如 28)个量级后，Ω_0 必十分地接近于 1. 这在当时是出乎宇宙学家意料之外的重要预言. 它对其他宇宙学问题的研究带来了深刻的影响. 即使对暴胀理论本身而言，有这种可供检验的预言是非常关键的事.

七、暴胀的启示及其产生机制

暴胀理论一出现人们就感到，它改变了我们对宇宙甚早期的理解，且这种改变是符合事实的需要的. 让我们对此作一回顾.

在没有暴胀的标准模型中所出现的疑难，其根本点是同一个：标准模型对经典宇宙提出了过分苛刻的初条件. 最原始的经典宇宙必须既超视界地均匀，又须在超视界范围内有显著超过随机涨落的密度起伏，此外初始膨胀速率还须与初始密度有精确的配合. 前面讨论过的正反物质不对称疑难也是同样的性质. 它要求在原初宇宙中正物质粒子就与反物质粒子有极微量的差别. 定性说来，经典宇宙无非是从量子宇宙转化而来. 人们感到不可思议：刚形成伊始的经典宇宙会满足如此苛刻的要求？

暴胀理论使人们对这些问题有了完全不同的理解. 它的要点也是一个：甚早期宇宙曾在一个短阶段里胀大了好几十个量级(例如 e^{100}倍). 前面已指出，经过暴胀，原来宇宙中的气体已被充分稀释，因此它对今天的宇宙状态已没有影响. 今天的宇宙介质是宇宙重加热的产物. 这样看来，原始经典宇宙的初条件失去了重要作用，宇宙重加热后的状态才是今天宇宙的实际初条件. 对于原初宇宙的研究，这无疑是一个崭新的认识. 对各种疑难的重新理解，则只是这新认识的自然后果.

宇宙学原理，即全宇宙的均匀性当然是一种理想化假设. 现在看来，整个经典宇宙有超视界的均匀是不可能的，但是存在一块很小的均匀区是可能的. 暴胀的效果把它变成一块很大的均匀区，以致我们的观测宇宙只是其中很微小的一部分. 这才是人们发现宇宙学原理，并能用观测证实它的基础.

以暴胀后为“初”，宇宙的视界已比观测宇宙小很多. 这时均匀性已经形成，而密度起伏必须也已形成. 所以相变中产生的涨落必须是结构形成的种子的来源. 关于原初正反物质的等量，原来是理

论家的猜想. 现在看来，宇宙重加热后的气体来自真空相变的潜热，它几乎必定是正反物质等量的. 因此，重子的产生必须曾发生过，且必须发生在暴胀之后.

最后不管原来宇宙密度多大，暴胀过程已把 Ω 对 1 的偏离压得十分低. 以暴胀后为"初"，尽管此后这偏离程度有几十个量级的放大，今天的 Ω_0 必定依然很接近于 1. 这些就是暴胀理论带给我们的启示.

暴胀并不是为解释某种疑问而设计的理论. 它是自然地使许多疑问得到了澄清. 这理论对宇宙学家的吸引力正在于此. 于是后继的研究迅速地发展起来了. 这里最重要的问题当然是：暴胀是否确实发生过？为澄清这一点，需要深入的是两方面. 一方面是关于暴胀的产生机制，即是否有可靠的物理理论，它保证暴胀一定曾发生. 另一方面是进一步研究暴胀理论的后果，看它们是否与观测事实相符. 本节先讨论前一方面.

最早的暴胀理论是把暴胀作为大统一破缺相变的后果. 虽然大统一理论本身是一种试探[①]，终究这是从其他物理领域为暴胀的发生提供了根据. 但是后继的研究很快认识到，从大统一相变来引发的暴胀是不符合宇宙学要求的. 于是人们把真空相变从大统一中分离了出来. 抽象地引入一种带自作用的标量场 ϕ，被称为暴胀子(inflaton). 适当地赋予它合适的性质，以使它产生符合宇宙学需要的暴胀. 这种研究进展得相当成功. 人们进一步想到，暴胀发生得很早，ϕ 场不应已有热平衡，面可能尚处于混沌状态. 混沌型的暴胀(chaotic inflation)模型是近年来许多很细致的研究的基础，其成功相当令人瞩目. 无论如何，有关的暴胀子是否客观存在是完全不清楚的.

有趣的是，对暴胀机制的广泛研究使人们认识到，可能在甚早期宇宙中诱发暴胀的物理机制很多. 现代版的 Kaluza-Klein 理论

---

① 现在粒子理论家已意识到，那种简单的大统一模型是不会成功的.

假定空间的维数远大于3. 额外的维数必须在宇宙演化中收缩到普朗克尺度，以使它看不到. 人们从这种研究发现，额外维的收缩也会诱发普通3维空间的暴胀. 此外有人指出，若引力常数在宇宙学时间尺度上有变化，它也能诱发暴胀. 这种类型的模型已有不少种，可是其中没有一种是具有被实验证实了的理论基础的.

在非常多样的研究中，还有一个结果很值得提到. 有人证明，要使视界疑难和准平坦性疑难等得到消除，宇宙甚早期曾有过一个短暂的加速膨胀阶段是必要条件. 如果广义地把加速膨胀叫暴胀（即不一定是指数型膨胀，也不一定由真空相变引起），这证明的意义是：若要不对经典宇宙的初条件提苛刻的要求，那么暴胀必须曾发生过.

在暴胀模型提出后的20年里，机制问题已被研究得很多. 我们只作如上极简单的概括，目的是说明事情有两方面. 一方面是至今没有理论上的论据说明暴胀确实发生过. 另一方面是人们有很多理由相信暴胀是曾经发生过的. 这就是问题的现状.

八、支持暴胀的证据

暴胀发生得很早，当时宇宙温度很高，因此它的理论基础是超高能物理. 现在的粒子加速器尚远没有达到这样的能量尺度，因此超高能物理一时还不会有明确的理论. 由于这原因，什么是暴胀的实际机制问题，是眼下不可能澄清的问题. 这样，暴胀理论的预言的检验就十分有价值.

要从这么早期的过程留下可观测的痕迹，这也是一件很困难的事. 实际上被认真对待的有两个问题：一是今天的宇宙密度；二是背景辐射场上的密度起伏.

暴胀理论预言今天的宇宙密度应等于1，或十分接近于1. 这是一个很有意思的预言. 原来天文学家已知道 Ω_0 大于0.2，但是不清楚会大多少. 如果暴胀理论是对的，这将意味着未被测到的暗

物质尚有很多倍. 那么,今天宇宙中不仅暗物质占绝大部分,而且绝大部分还不是普通物质(参看第十章). 因此要是能肯定这是事实,不仅对暴胀理论是重要支持,而且对其他宇宙学意义也很大.

80 年代中期,对宇宙密度的全局性测量有了几个尝试. 在较大的统计误差下,得到了支持 Ω_0 接近于 1 的结果(再参看第十章). 但是争议很大. 问题还不仅是统计误差,人们对这方法的可靠性不信任,即认为它的系统误差也很大. 因此问题的答案仍不明确. 一直到 1999 年,有人从微波背景上测到的扰动谱推出 Ω_0 很接近于 1. 这结果的争议较少. 因此我们可以期望,在今后 10 年里,这问题的答案可能将得到澄清.

此外最受人注意的,就是暴胀引起的原初扰动在背景辐射上留下的痕迹. 从 1992 年起,背景辐射上的密度起伏已被观测到. 我们也已知道,若没有暴胀,那么人们原则上不能追究它的起源. 按暴胀模型,它应是暴胀过程留下的遗迹. 仅从这意义上,暴胀的作用就够重要的,但是人们想有更强有力的证据.

这问题上要做的事又分两头. 至今对背景辐射上的扰动的测量还远不够细致. 有两颗科学卫星将在近年内发射上天. 一颗叫 MAP,另一颗叫普朗克巡天器. 按计划,它们将极细致地测量出扰动的多极分布. 它的结果将对宇宙学至关重要. 问题的另一头是用暴胀理论计算应有的扰动大小. 尽管理论结果会依赖模型细节,无论如何,有了观测结果对人们了解这些扰动是否是暴胀的遗迹将很有帮助.

总之,现有的检验已显示出支持暴胀理论的迹象,我们能够期望在不久的将来,看到更强烈的支持证据.

第九章　宇宙常数与真空能

宇宙常数原是爱因斯坦出于一个错误的动机而引入的.他后来倾向于取消它.但是量子场论的发展表明,若真空有能量,宇宙常数是不能取消的.现今令人困惑的是,实测已知道宇宙常数很小,而理论估算出的值却比它大几十个量级.两者完全不能协调.这有可能暗示了人们的真空概念有根本性的错误.如若如此,这将是物理学发展的契机.

一、问题的历史

我们已经知道,爱因斯坦于1915年建立的引力场方程中是不包含宇宙常数项的.当他用它研究宇宙时,发现只有引力存在使宇宙不可能静止,而他却先验地相信宇宙应是静止的.为了做成一个静态的宇宙模型,他在引力场方程中加上了宇宙常数项.宇宙项提供的是斥力.有斥力与引力相抗衡,宇宙才能够静止.爱因斯坦最初就是这么把宇宙常数引进物理学的.当时是1917年.

不几年后,天文学家开始发现了宇宙在膨胀的初步迹象(参看第一章).爱因斯坦敏感地接受了这事实的启示.他在与朋友Weyl的通信中说,如果不存在接近静态的宇宙,那么把宇宙常数去掉吧.那时是1923年.对于引力场方程中是否应包含宇宙项,在不到10年内就出现了一次反复.其实这只说明当时对这问题的认识还很肤浅.

归根到底讲，要或不要宇宙常数，这是只能由实测回答的问题. 后来它变成一个深刻的物理问题，这是微观物理发展的后果. 1926 年，狄拉克建立了相对论性的电子运动理论. 今已周知，狄拉克方程有负能解，使电子运动失去了基态. 于是狄拉克提出见解：一切负能态已被充满时才是真空. 人们把这样的真空叫狄拉克的电子海. Gamow 立即意识到这样不行. 这样的真空（即电子海）将有很大的密度，它是会产生很强的引力的.

此后不几年，量子场论出现了. 在量子场论中，负能级自然地描写了反电子的行为. 狄拉克把电子海当真空的概念也就不需要了. 但是按量子场论，真空是否也有密度？这新问题又出现了. 有趣的是，写出这真空的能量-动量张量，它与宇宙常数项完全一样. 这一来，宇宙常数是否为零与真空能密度是否为零，变成了两个分不开的问题. 由此，一般引力场方程中要不要保留宇宙常数项，成了一个深刻的物理问题.

二、真空有能量吗？

这问题涉及到两个很不简单的物理概念：什么是能量和什么是真空. 让我们从讨论能量概念入手.

能量概念原来是依赖于它的守恒性而存在的. 如果把能量守恒理解为运动不生不灭的体现，那么能量就是对各种不同形态的运动的量度. 物理实践已表明，一个物体能量的减少，必有另一个物体能量的增加；一种运动形态（力学的、热的或电磁的）能量的减少，必伴随有另一种运动形态的能量增加. 这就是运动的守恒性. 从这一意义上讲，能量的定义中只有可变化的部分才有物理意义. 若在其上叠加一个任意常数，它对能量的守恒性将没有影响. 人们熟知的例子是重力势能. 其实动能的定义也能差一个常数. 之所以没有这么做，仅是出于自然性，即让不动物体的动能为零. 这并不是必须的.

狭义相对论中的质能关系给概念带来了复杂性. 在自然单位制下,质能关系表达为 $E=m$. 这样更清楚地示出,质量和能量虽然是两个不同的概念,但是它们的大小由同一个物理量描述.

质量本身原是一个双重性的概念. 牛顿第二定律中的质量描述运动物体的惯性. 牛顿引力定律中的质量描述物体引力作用的强弱. 牛顿曾思考过,为什么这两个不同的物理概念会用同一个量来描写. 他没有能作理论上的回答,但是用单摆实验证明了这是客观事实. 因此后来的物理学中不再区分描述惯性的质量 $m_{惯}$ 和描述引力的质量 $m_{引}$. 让一个物理量 m 描述物体的两重性质.

在这基础上,质能关系又进一步表明,质量或能量是同一个物理量. 这样这个量描述了物体的三重属性: 运动的剧烈程度、引力"荷"(指引力质量)的大小和惯性的大小. 注意引力质量及惯性的大小是有客观效果的. 其值越大,所产生的引力将越强,及不易被加速的程度也越大. 它们的定义完全不允许叠加一个常数. 这样能量(即质量)也就不再有零点的任意性了. 这是要讨论真空能之前须澄清的一个要点.

然后考虑什么是真空. 经典物理中的真空就是没有物质的状态. 这概念较简单. 这样的真空不具有能量也几乎无可争议. 问题的复杂性来自微观物理.

从量子场论看,一切微观物质都以量子场的形式存在. 当场受到激发,我们才"看到"了场量子,质子和电子等都是相应场的量子. 若场处于最低能态(叫基态),就是没有相应的粒子. 因此一切量子场都处于基态时,我们看不到任何粒子. 理论上就把这一状态叫真空态. 它是宏观真空概念的自然延伸,但值得注意其差别. 这样的真空中没有粒子,而处于基态的量子场本身依然存在. 真空已不是空无一物的状态了.

量子场论中的真空即基态,它是个复杂的物理状态. 例如真空中会有虚粒子对的产生和湮灭,叫真空涨落. 在量子电动力学中算出过真空涨落引起的可观测效应. 它与事实符合得很好. 这说明真

空确非空无一物，而仍有着丰富的物理内涵. 因此真空是否具有能量，就成了一个合理的物理问题.

为说明量子系统处于基态时有能量，谐振子是很好的例子. 量子力学告诉我们，任一频率的谐振子处于基态时有能量$\frac{1}{2}h\nu$. 基态作为最低能态，虽然已没有能量可以输出，但它依然是具有能量的. 这其实是测不准原理的直接后果. 要一个微观粒子静止地处于势能最低点是不可能的. 量子场可用傅里叶分析的观点，看成不同频率的谐振子的叠加. 那么当量子场处于基态，它也会是有能量的.

按上面的分析，若真空有能量，它将有三种物理效果. 惯性效果在这里没有意义，因为人们不能让真空加速. 此外在粒子物理现象中，引力效果小得无法观测，因而可被忽略. 这样能量只剩下一重原始的含义，即在转化中保持守恒. 在这意义上，真空能已不能转化，因此其大小将不起作用. 这就是能量零点的任意性. 正是出于这道理，在作量子场论计算时，人们可把真空能密度规定为零.

要是真空实际上有能量，在能量守恒和在有惯性两方面是反映不出来的，所以真空是否有能量的问题仅在引力问题中才有意义. 若它有能量，真空将也会产生引力(下面将看到实际是斥力). 人为地把它规定为零是不行的. 这正是当年 Gamow 从狄拉克海中感到的问题. 现在宇宙学关心真空能，也是出于同样的想法：若真空有能量，它产生的引力会对宇宙膨胀有影响的，而且这影响是可观测到的.

三、真空产生的引力

真空的引力效应只能在广义相对论的基础上研究. 为此把真空当引力源的一部分，先得写出它的能量-动量张量$T_{\mu\nu}$. 基于原则上的考虑，这能量-动量张量只可能是度规$g_{\mu\nu}$乘上一个常数. 我们

把它写成

$$(T_{\mu\nu})_{vac} = -\rho_{vac}g_{\mu\nu}. \tag{9.1}$$

在这样的形式下，ρ_{vac}代表真空的能量密度.

把真空的贡献和宇宙常数的影响都考虑在内，引力场方程应重写作

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R - \lambda g_{\mu\nu} = -8\pi G(T_{\mu\nu} - \rho_{vac}g_{\mu\nu}). \tag{9.2}$$

这里的$T_{\mu\nu}$理解为常规的、不包括真空贡献的能量-动量张量. 写出这一般方程立即看出，宇宙常数对方程的影响与真空能密度一样，都是添加一个度规乘常数的项. 这两项总可以合并，说明两者的影响是分不开的.

有两种形式合并它们. 一是引入等效宇宙常数

$$\lambda_{eff} = \lambda + 8\pi G\rho_{vac}. \tag{9.3}$$

把方程(9.2)化成

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R - \lambda_{eff}g_{\mu\nu} = -8\pi GT_{\mu\nu}. \tag{9.4}$$

右边不再有真空能项，但其效果已包括在内了. 另一个形式是引入等效真空能密度

$$\rho_{eff} = \rho_{vac} + \lambda/8\pi G, \tag{9.5}$$

把方程(9.2)写作

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = -8\pi G(T_{\mu\nu} - \rho_{eff}g_{\mu\nu}). \tag{9.6}$$

方程左边不再有宇宙项，但其效果也已在内了. 很显然，这三种形式都等价.

把这方程用于均匀各向同性的宇宙，导出动力学方程是

$$\dot{R}^2 + k = \frac{8\pi G}{3}(\rho + \rho_{eff})R^2, \tag{9.7}$$

$$\ddot{R} = -\frac{4\pi G}{3}(\rho - 2\rho_{eff})R. \tag{9.8}$$

它就是第三章中讨论含宇宙常数模型的出发方程(参看(3.1)和

(3.2)).差别只有一点.那里 ρ_{eff}是宇宙常数的代表(见(3.3)),而现在它是宇宙常数和真空能的总效果(见(9.5)).这样看来,考虑到真空有能量,宇宙的动力学行为不需重新研究.它已在含宇宙常数的模型中讨论过了.

无论如何,这里出现了一个物理学中前所未有的奇怪境地.如果引力场方程中应包含宇宙项,那么宇宙常数 λ 就和牛顿引力常数 G 一样,是一个基本物理常数.它的值只能从实验来定出,而不可能由理论推断.上面的分析使我们看到,实验只能定出等效宇宙常数 λ_{eff}或等效真空能 ρ_{eff},它是 λ 和 ρ_{vac}的联合效果.这意味着宇宙常数 λ 本身既不能由理论定出,也不能由实测定出.这真是物理学的尴尬.

幸好真空能密度 ρ_{vac}不是基本物理常数.它作为派生的常数,原则上可以由量子场论算出.这样总算有一条出路:用实验测出等效常数 ρ_{eff}(即 λ_{eff}),再从理论上算出 ρ_{vac},宇宙常数 λ 就能由(9.5)(或(9.3))定出.

爱因斯坦倾向于认为 $\lambda=0$.如果他是对的,那么测出的 ρ_{eff}和算出的 ρ_{vac}应该相等,否则就是他作出了错误的判断.后人终究是能把这问题澄清的.可是宇宙常数问题的研究竟然是如此的戏剧性:理论与实测对比的结果,出现的是很大的困惑.

四、问题的困境

问题的实测方面已在第三章中讨论过.人们在十几年前定出的上限是

$$\rho_{eff} < 0.9\rho_c. \tag{9.9}$$

这结果并不排除 $\rho_{eff}=0$.近年定出的值是

$$\rho_{eff} = 0.65\rho_c. \tag{9.10}$$

总之,若 ρ_{eff}不为零,它的值总是略小于临界密度 ρ_c.下面用三种单位写出 ρ_c 的量级,以便对它有感性的了解:

$$\rho_c = \begin{cases} 10^{-46}\,\mathrm{GeV}^4 & (\text{自然单位}), \\ 10^{-4}\,\mathrm{eV/cm}^3 & (\text{能量密度单位}), \\ 10^{-29}\,\mathrm{g/cm}^3 & (\text{质量密度单位}). \end{cases} \qquad (9.11)$$

与任何非宇宙学的环境相比，这是一个很小的密度.

然后转向问题的理论方面，即作真空能密度的估算.

原则上须知道一切基本量子场的形式，以及其间的一切相互作用，才有可能计算真空的能量密度. 这不是目前能做到的事. 无论如何，能量有可加性，我们可以先对其中一部分做估算. 这样可期望知道它的量级.

先最简单地估算一种质量为 m 的自由标量场的基态能密度. 对场做傅里叶展开，它就成了一切频率的谐振子的叠加. 把每一频率的谐振子的基态能相加，得到

$$\rho_{\text{vac}} = \frac{1}{(2\pi)^3}\int_0^{\Lambda} \frac{1}{2}\sqrt{k^2 + m^2}\cdot 4\pi k^2 \mathrm{d}k \approx \frac{\Lambda^4}{16\pi^2}, \qquad (9.12)$$

其中 Λ 代表频率的上限. 若频率没有上限，其基态能密度将是无穷大. 有理由认为对超过普朗克能量的粒子将不能这样估算. 把上限取作 $\Lambda = 10^{19}\,\mathrm{GeV}$，估出

$$\rho_{\text{vac}} = 10^{74}\,\mathrm{GeV}^4. \qquad (9.13)$$

这真空能密度竟比实际的 ρ_{eff} 高 120 个量级以上！让我们看这意味着什么.

宇宙常数 λ 应由(9.5)定出. 从中看到，$\frac{\lambda}{8\pi G}$ 须与 ρ_{vac} 十分接近，并取反号. 观测到的 ρ_{eff} 是两者前 120 位相抵消后极微小的余量. 这样意味着要谈 λ 的大小，必须有 120 多位有效数字的精度. 若在前 120 位中改变一点，抵消后的余量 ρ_{eff} 将增大许多量级. 注意这余量才会对宇宙有实际的影响. 于是这增大了很多量级的 ρ_{eff} 的斥力将使宇宙转瞬间就从高密状态变成极稀薄状态. 这样的宇宙中根本来不及形成星系，更不会有人类. 理论上的宇宙与实际将完全不一样. 于是我们只能得出结论：今天的宇宙之所以是看到的那样纯属偶然，它是宇宙常数取值在 100 多位有效数字上细微

调节的结果.若它变一点,今天宇宙将完全是另一个样子了.

这样的结论是物理学家难以相信的.如果说原来只是对宇宙常数还不够了解,进展到这地步,问题却是遇到了真正的困境.

五、理论上解困的失败

20世纪70年代中期,粒子理论出现了超对称的想法.按这理论,自然界存在着费米子与玻色子之间的对称.任何一种费米子都有同质量的玻色子相对应,反之也一样.但是已知粒子如光子(玻色子)、电子(费米子)等都没有发现其对应粒子存在.这表明如果超对称理论是对的,它也只能是在低能下破缺的对称.

随后有人证明了一个很有启发性的结果.若超对称严格成立,那么同质量的费米场和玻色场的真空能将抵消,即它们的总真空能密度是零.这样一来,真空能太大的困境就没有了.但是现实的超对称不可能严格成立,它必须在低能下破缺,否则与事实显然不符.因此他所讨论的不是真实模型,而只是玩具模型.无论如何,理论家从这玩具模型中感受到了一种启示.

困境的出现是由于我们只计及一种场的真空能.可能高能物理中存在着新的原理(例如上述的严格超对称),它保证着一切场的总真空能密度为零.要是如此,我们就摆脱困境了.许多粒子理论家按这思路作了大量的尝试.

由于超对称的破缺,对应的费米子和玻色子在低能下会有显著不同的质量.我们没有发现光子和电子等一切粒子的超对称对应物,正是这原因.从这样的事实看来,超对称的破缺能量 $E_{超}$ 至少在 10^2 GeV 以上.更高能量下的真空能会因超对称成立而抵消.估出低能下不完全抵消所留下的真空能密度,它依然很大.达到

$$\rho_{vac} \approx E_{超}^4 \geqslant 10^8\ \mathrm{GeV}^4. \qquad (9.13)$$

这能量密度比 ρ_c 还高出54个量级.因此,有破缺的超对称理论完全没有使宇宙常数问题得到缓解.

把上述超对称性质局域化，引伸出来的是超引力理论. 人们自然想继续探索：能否以此为基础证明总真空能为零. 结果收获甚小. 先是有人用特殊模型证明了，考虑到超对称的破缺，真空能的最低级微扰会抵消为零. 这是好的苗头，但是后续的进展迟迟出现不了. 最低级微扰下真空能为零的结果没有得到普遍的证明. 高级微扰下的真空能是否会逐级抵消，更不清楚. 此外有人指出，即使每一级微扰都抵消，非微扰效应还会留下很大的真空能. 总之，人们没有从超引力理论中看到解困的希望.

粒子理论中接着出现的新概念是超弦. 同样目标的研究在继续，但是没有惊人的喜讯. 让我们用著名粒子理论家 Weinberg 的话来结束这段讨论. 他说："我们难以从超引力或超弦的任何性质中看到使等效宇宙常数充分地小的可能."他说这话的时间是 1989 年，但至今情况没有实质性的变化.

考虑到上一章讨论过的真空相变，它只把问题推向更大的困境. 上面谈论的是动力学真空，即零温真空. 从宇宙学的角度看，它是超高温真空经过逐级相变而降下来的. 大统一相变($T\approx10^{15}$ GeV)使真空能密度下降了 $\Delta\rho_{vac}\approx10^{60}$ GeV^4. 弱电破缺相变($T\approx10^2$ GeV)又下降了 $\Delta\rho_{vac}\approx10^8$ GeV^4. 接着的夸克-强子相变($T\approx0.5$ GeV)再下降 $\Delta\rho_{vac}\approx0.1$ GeV^4. 我们需要的是经过这些相变后，今天的零温真空有

$$\rho_{vac}\approx\rho_{eff}<10^{-46}\ \mathrm{GeV}^4. \qquad (9.14)$$

这又是一个很不自然的要求.

总之，今天的零温真空能密度实际上很小，而理论上估算得到的值总是很大，这令人困惑的局面至今没有看到出路.

六、困境的启示

在宇宙常数问题上，理论上看来应当有的结果却与事实明确地不一致，这使人联想起 100 年前物理学上出现过的情景. 当时有

两个类似的矛盾，被开尔文称为物理学上空的两朵乌云．后来都变成了物理学中耀眼的星光．借鉴一下历史是有益的．

两朵乌云之一是迈克耳孙-莫雷实验的零结果．当时日心说与地心说之争早已结束．人们知道地球不会绝对静止，这是肯定的．再进一步，因地球以 30 km/s 的速度绕太阳转动，那么不管太阳怎么运动，地球的绝对速度在一年中总会有时候要超过 30 km/s．这几乎也是无可争议的．然而，迈克耳孙-莫雷经过几年的观测后说，地球对以太的速度一定小于 5 km/s，否则我们不会老是得到零结果．这就是事实与当时公认的理论推断出现矛盾的例子．

物理学是一门实验科学．当有这种矛盾出现，原则上错误总在理论一方（实验不可靠除外）．理论推断被公认，只说明错误出在某个被大家习以为常的概念上．对于这个例子，今天已周知，需要的是抛弃以太和改变时空观念．这种基本概念上的变化，给物理学带来的影响常很深刻．这朵乌云的消散，导致了相对论的建立．它后来成了 20 世纪物理学的理论基石之一．

另一朵乌云的本质很类似．如果经典统计对辐射场适用，那么热辐射的紫外灾难是不可避免的．由于能量按自由度的均分，结果注定如此．但是观测事实却清楚地表明，频率越高，热辐射越弱．在当时，理论与事实两方面都几乎无可指责．然而它们却完全不一致．这才是深刻的矛盾（或称危机）！现在我们都知道，根源在于辐射场有量子性．可是物理量的连续性是大家习以为常的观念，当时的人们怎么会想到问题出在这里！？无论如何，这朵乌云的消失，成了量子物理兴起的契机．量子理论后来是 20 世纪物理学的又一重要基石．

正是历史的借鉴使理论家 Weinberg 感到，危机使物理学兴旺[①]，可惜后来出现的危机太少了．有重大的物理问题没有解决，或无法解释并不算危机．这是永远会有的事．重要的是看来不可避

① 原话是：Physics thrives on crisis．

免的理论推断与肯定了的事实出现冲突．这才值得被称为危机．的确，这样的危机并不经常出现．但是它的出现，多半表明人们某种习以为常的观念有错误．

回到宇宙常数和真空能的问题上来．现在人们所困惑的正像是这种冲突．不管近年对等效真空能的测定是否可靠，真空能密度不能太大，这已是无可怀疑的事实．理论上，经过近 30 年的广泛研究，人们发现真空能似乎不可避免地很大．这构成了物理学史上难得出现的冲突．它使我们思想上又回到最初的问题：到底什么是真空？它是否有能量？这无疑是一个有趣而深刻的问题．

有人开玩笑地说，亚里士多德讲过自然忌真空，现在要讲的或许是其反面：自然忌真空不空．当年电磁理论把以太塞进真空，大自然不接受，用相对论把以太赶出了物理学．现在的量子场论好像又在真空中塞进了过多的物质．它最终会在某种形式下被赶出去的．

第十章　暗物质问题

宇宙密度多大?物质的主要组分是什么?这是宇宙学的基本问题.但是这样的问题竟然很难回答.首要原因在于人们主要凭借光来作观测,而宇宙中绝大部分是不发光的暗物质,所以只能靠它们的引力效果来推断.现今的结果很使人惊讶.宇宙中最主要的组分是真空,其次是不会发光的非重子气体.由原子组成的普通物质肯定只占百分之十以下.这三种物质合在一起,总密度大约正好与临界密度一样.

一、星系的暗晕

宇宙很可能是无限的,所以它的物质总含量只能用密度来描写.这是一个必须由实测确定的基本参量.确定密度就是要测量单位体积内的质量.把宇宙气体看成由星系所组成,那么我们须先会测量星系的质量.

类似于我们银河系的旋涡星系是靠转动维持平衡的.距中心为 r 处的恒星具有的转动速率 v 应满足

$$v^2 = GM(r)/r, \tag{10.1}$$

这里的 $M(r)$是半径为 r 的球内的质量.天文学家能测量星系的转动曲线,即测量 v 作为 r 的函数.这样就能把星系内的质量分布确定下来.

星系内的恒星分布区有明显的边沿.边沿之外恒星已很少,但

稀薄气体是依然存在的. 就像地球的固体表面之外有大气一样. 原来以为星系发光区外气体的总质量很次要. 到 70 年代, 人们才发现, 那些气体的总质量比全部恒星的总质量大若干倍.

星系发光区外的物质叫暗晕. 暗晕中的氢原子 21 cm 谱线是可观测的. 通过测量这谱线的多普勒频移, 可以推出该处物质对星系中心的转动速率. 于是实测的星系转动曲线可延伸到发光区之外. 如果星系的质量主要集中在发光区内, 那么外部的 v^2 应当与 r 成反比地下降(见(10.1)式). 可是对大量旋涡星系转动曲线的测量表明, v 几乎不随 r 增大而下降. 这情况被理想化后画在图 10.1 中.

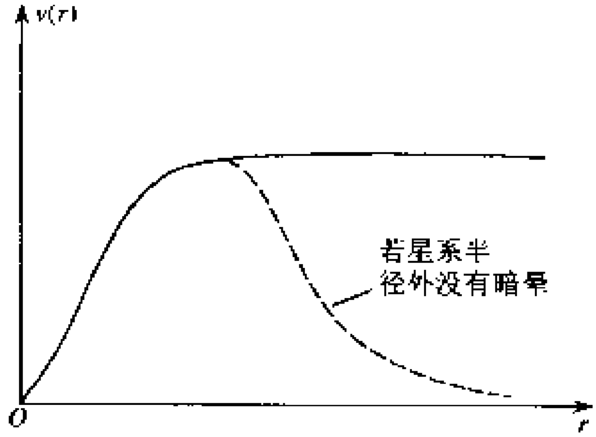


图 10.1　星系转动曲线的平坦延伸

从关系式(10.1)看, 转动曲线平坦延伸, 表明 $M(r)$ 与 r 成正比. 实际对不同星系测到的转动曲线平坦延伸至发光区半径的 3 至 10 倍间. 这意味着把暗晕物质考虑在内, 星系的总质量比只计及发光部分时要增大 3 至 10 倍! 这是星系中不发光物质为主要部分的强烈证据.

转动曲线的平坦延伸是否说明气体的密度分布很反常? 其实不是. 容易论证, $M(r)$ 正比于 r, 表明气体密度 ρ 与 r^2 成反比地降低. 越外边的气体还是显著地越稀薄. 事实上人们对转动曲线平坦延伸的范围不清楚, 就是因为气体太稀薄处的氢 21 cm 谱线已测不到了. 上面提到的 3 至 10 倍处, 并不是暗晕的边缘.

让我们注意，天文学家平时讲的星系质量，指的是其发光区内的总质量. 连暗晕在内，星系的质量大小不大清楚. 上述测定和讨论说明，实际总质量估计可能要大 3 至 10 倍. 旋涡星系如此，椭球星系也如此.

二、星系团内的暗物质

星系团可看成许多星系在自引力作用下的平衡系统. 设 r_i 为第 i 个星系与系统质心间的距离，则各个距离平方的加权平均值是

$$R^2 = \sum m_i r_i^2 / M, \tag{10.2}$$

其中 m_i 是第 i 个星系的质量，$M = \sum m_i$ 是系统的总质量. 统计物理中叫这 R 为系统的维里半径. 当星系在相互引力作用下运动，而其维里半径保持不变，则叫系统处于维里平衡. 这时的 R 可作为星系团的大小. 由牛顿第二定律可证明，维里平衡下的 R 满足

$$v^2 = GM/R, \tag{10.3}$$

其中 v^2 是各星系速率平方的加权平均. 从(10.3)看，测量星系团的大小，以及成员星系的无序程度，就能推知星系团的总质量.

天文学家 Zwicky 在 1933 年就用这方法定出了星系团的质量. 他意外地发现，所定出的质量比其成员星系的质量和大 10 至 100 倍. 这是第一次发现暗物质的总质量成倍地比发光物质多的事例，但是当时人们并不认为这是需要认真对待的结果. 在星系暗晕中包含着很大质量的事实得到确认后，星系团中的暗物质问题才开始有系统的研究.

现今对星系团的测定，发现它的质量占其成员星系总质量的 10 至 30 倍，即当年 Zwicky 的结果是大体可靠的. 这里的星系质量仍仅指发光区内的. 考虑到暗晕也是星系的一部分，估计出星系团层次上的暗物质还比其中星系的总质量大几倍. 这新层次上的

暗物质主要是星系际气体.由于星系际的空间容积很大,包含了如此大量物质的气体依然是十分稀薄的.它的密度仅比宇宙平均密度大一个量级左右.

这些从实测得到的结果使我们对宇宙介质的认识深化了一步.在整个宇宙中,发光的天体仅是一小部分,不发光的物质才是主要部分.像发光天体有层次性结构一样,暗物质的分布也是层次性的,星系内和星系团内存在着各自的暗物质.

三、宇宙密度的局域测定

我们简化地把宇宙介质看成由星系为"分子"的均匀气体,那么其密度 ρ 是单位体积内的星系个数 n_G 与星系平均质量 M 的乘积,即

$$\rho = n_G M. \tag{10.4}$$

不同星系的质量差别很大.巨型星系的质量有 $10^{12}\ M_\odot$,而矮星系的质量仅为 $10^6\ M_\odot$.矮星系的数量很多,而巨星系的数量要少得多.因此星系的平均质量极难可靠地确定.

考虑到大星系必很亮,而小星系必很暗,不同星系的质量和光度之比 M/L 不会有太大的差别.于是把上述密度公式改为

$$\rho = (n_G L) M/L = \mathscr{L} M/L, \tag{10.5}$$

其中 $\mathscr{L} = n_G L$ 是宇宙中的光度密度,也就是单位体积内星系的总光度.这光度密度能够直接测量.它的值是

$$\mathscr{L} = 1.2 \times 10^8\ L_\odot/\mathrm{Mpc}^3, \tag{10.6}$$

其中 $L_\odot$ 代表太阳光度,它被用作光度的单位.这样,我们把宇宙密度的测定转化成了局域的测量:测量不同星系的平均质光比 M/L(以太阳的质光比 $M_\odot/L_\odot$ 为单位).测得后把它与由(10.6)表示的光度密度相乘,得到的就是宇宙密度(以 $M_\odot/\mathrm{Mpc}^3$ 为单位).

星系的光度 L 较易于直接测定,问题回到星系的质量 M.现在让我们分层次讨论.若只计及星系发光区的质量,不同星系的质

光比为

$$(M/L)_{光} = 3-10. \tag{10.7}$$

这意思是说，星系中每发出 1 L 的光，背后有 3 到 10 倍太阳质量的物质. 把这结果代入，即可定出宇宙中发光物质[①]对密度的贡献. 用无量纲密度 Ω 表示，它是

$$\Omega_{光} \approx 0.01. \tag{10.8}$$

上面已讨论过，实际上暗晕才是星系中的主要物质. 把它考虑在内，星系的光度没有增加，而质量要增大 3 至 10 倍. 把星系质光比加大 3 至 10 倍后，推出一切星系对密度的贡献为

$$\Omega_{星系} \approx 0.03-0.1. \tag{10.9}$$

再注意到星系际气体的质量更大，它对宇宙密度的贡献必须算进去. 这样就要测定星系集团的质光比. 它的光度来自各星系，要点仍在于其总质量的测定. 以较小的星系集团，即星系群做单元，得出的质光比是

$$(M/L)_{群} = 20-70. \tag{10.10}$$

以富星系团为单元，其质光比是

$$(M/L)_{团} = 100-300. \tag{10.11}$$

当采用局域的测量定总体密度，当然单元取得越大，包含进去的暗物质越多. 用大星系团的质光比代入，得到密度为

$$\Omega_{团} \approx 0.1-0.3. \tag{10.12}$$

以超团为单元做估计，结果是差不多的. 考虑到以星系集团为单元做局域测量时，各单元间存在的暗物质总没有包括在内. 所以我们只能得到结论：宇宙总密度是

$$\Omega_0 \geqslant 0.2 \pm 0.1. \tag{10.13}$$

把结果写成不等式意指，已测到的暗物质肯定是存在的，而未测进去的更弥漫的暗物质未必是可以忽略的.

---

① 这里指一切星系发光区内的物质总和.

四、密度作为宇宙学参量

在地面上测量物体的质量，多半是利用它受到的引力（即重力），而测量天体的质量却只能利用它所产生的引力. 按照牛顿引力定律，对球对称的系统，半径为 r 的区域外的物质在区域内的引力相互抵消了. 因此通过星系内或星系团内的运动测量，不能得到外部存在的引力源的质量. 这就是上面讨论的局域测量的局限性. 可是天文观测已证实，星系或星系团外的物质虽然稀薄，但是其总质量并不小. 这样，为了确定宇宙物质的总量（用总密度描写），我们必须测量宇宙的整体运动，因为它才反映了宇宙中全部物质的引力.

宇宙整体运动满足 Friedmann 方程. 为把运动完全确定，还须把今天的密度 Ω_0，膨胀速率 H_0 和宇宙常数作为参量输入. 对宇宙做整体测量能把宇宙密度定下来，其道理就在这里. 整体性测量的种类很多. 远处星系的红移-距离关系的测定是其中之一. 此外较被常用的是星系计数与红移关系的测量.

后一方法所测的，实际上是共动体元[1]的体积随距离的变化. 越远的距离上看到的是越早的历史，所以计数-红移关系反映的是宇宙膨胀的历史. 我们因此也能用它来确定宇宙的密度. 使用这样的方法，要求对一个天区（即一个立体角）内的所有星系作红移测定. 红移的大小是空间远近的标志，也是历史早晚的标志. 为把宇宙的"历史"反映出来，所取的红移限还不能太小. 所以这测量实际是很难做的.

按照标准模型，单位红移间隔内的星系个数 dN_G/dz 随红移 z 的变化规律应是

---

① 共动体元指一块固定物质构成的体元.

$$\frac{dN_G}{dz} \propto \frac{[\Omega_0 z + (\Omega_0 - 2)(\sqrt{1+\Omega_0 z} - 1)]^2}{H_0^3(1+z)^3\Omega_0^4\sqrt{1+\Omega_0 z}}, \quad (10.14)$$

其中 Ω_0 和 H_0 作为参量出现. 在实测和理论的拟合中,可以把密度定出. 各种全局性的密度测量方法都是同样的精神,也都有类似的困难.

80 年代中期,曾有一个研究组用 $z<0.75$ 的 1000 个场星系①为统计样品,研究了星系计数-红移关系,他们在 95%的置信度上定出

$$\Omega_0 = 0.9^{+0.7}_{-0.5}. \quad (10.15)$$

这结果第一次较认真地表明宇宙密度很接近与 1,从而表明在星系团尺度之外还有相当大量的暗物质. 这结果的意义当然很重大. 更使理论家感兴趣的是它支持了暴胀理论的预言. 但是这测定确实有很多可争议之点.

上述理论公式中假定了从那时(例如红移为 0.75 时)到现在,星系既没有产生也没有死亡. 当时认为,若涉及的历史不长(例如 $z<0.75$),实际情况偏离这假定不会太远. 可是后来有观测证据表明,有些矮星系形成得很晚(例如在 $z=0.2$ 前后). 更大的问题来自观测资料的不完备. 红移 $z=0.75$ 意味着光度距离达到了 10^3 Mpc. 那里的大星系被看到了,而同样距离上的小星系却可能没有看到. 这使统计资料不符合理论公式的要求. 此外还有其他疑点. 总之,这结果只能看成是很初步的.

用距离-红移关系定密度,必须涉及 $z\geqslant1$ 的星系的测距. 由于远距离定标的困难,长期来没有可信赖的结果. 后来因在距离定标上找到了较好的办法,才使 Perlmutter 等人在 1998 年得到了较好的结果. 这已在第三章中讨论过了. 近年来,这项结果和 BOOMERanG 对背景辐射上的扰动测定已使很多人相信,宇宙密

① 场星系指不属于大型星系集团的星系.

度 Ω_0 确实很接近于 1. 我们有理由预料，这问题在十年内将会有无可争议的回答.

五、非重子暗物质的概念

宇宙中有大量暗物质存在，这不是一个难以想像的结果. 按天文学知识，天上肯定有很多类型的天体是不发光①的. 让我们想得具体一些.

在星系之内，除看得见的恒星外肯定有看不见的恒星. 质量显著比太阳小，光度也显著比太阳小的叫矮星. 它们稍远一些就看不到了. 若质量小于 0.08 $M_\odot$，理论上知道它内部不能引起氢的燃烧，因而形不成恒星. 它们应归入行星的一类. 它们也将表现为不发光的暗物质. 恒星演化到晚期，将变成白矮星、中子星或黑洞. 这些形态的恒星内部已失去了能源，因此发光的同时在冷却，并逐渐变得越来越暗. 它们也是暗物质的一种形式. 此外，恒星际有稀薄气体，它也是不太发光的物质.

在星系团内，主要能想到的暗物质是矮星系和星系际气体. 事实上，有大量的矮星系（$M\sim10^6\,M_\odot$）的存在是在后来才知道的. 星系际的热气体会有 X 射线辐射. 它表明其总质量不少，也是近十来年才确证的事.

对更大尺度上的暗物质，人们易于想到的是稀薄气体. 此外还可能有不发光的星系. 星系的发光靠内部恒星的大量形成. 若某些星系质量的团块不具有在内部形成大量恒星的条件，它们将是看不到的. 这种团块被叫作未成熟的星系（failed galaxy）.

暗物质很多并不意外，因此它并不会引起宇宙中有非常规物质存在的联想. 非重子暗物质概念的起因很有些偶然. 但是它出现

---

① 不发光的说法很模糊. 若把一切波段的辐射都算在内，天体或多或少地都发"光". 这里让我们理解为：很多类型天体的辐射很弱，在宇宙总辐射中很次要.

后立即显现了很强的生命力，以致非重子暗物质很快就成了宇宙学理论研究中不可缺少的概念.

80 年代初，一位苏联的物理学家 Liubimov 发现(与电子同代的)中微子有静质量，而且高达 30 eV. 这结果在粒子物理学界引起的是怀疑，而在宇宙学界引起的却是兴奋. 宇宙学家早已知道，今天宇宙中不仅有光子背景，而且有中微子背景. 它们都是从早期宇宙中退耦后遗留下来的. 它们的粒子数密度都很大. 背景中微子约在每立方厘米有 100 个. 如果每个有 30 eV 的静质量，那么宇宙密度 Ω_0 将主要来自中微子，而且单凭它的贡献，Ω_0 就显著地大于 1. 这当然是一个很令宇宙学家兴奋的新结果.

很快 Liubimov 的静质量测定遭到了理论和实验方面的否定. 但是宇宙中以非重子为主的概念却被保留了下来. 理由是多重的.

几乎同时出现的暴胀理论预言了 $\Omega_0=1$. 此后不几年，原初核合成理论的检验预言了重子物质的总量为 $\Omega_B=0.011-0.19$ (1984 年). 按近年的结果，较肯定地有

$$\Omega_B < 0.10. \tag{10.16}$$

这两个预言结合起来，宇宙中 90%以上的物质不是由质子和中子等粒子组成的. 这是使理论家相信宇宙以非重子物质为主的重要根据. 如果考虑到当时 $\Omega_0=1$ 尚没有证据，那么从密度局域测量得到的 $\Omega_0\geqslant 0.2$ 也强烈地暗示了非重子是暗物质的主要组分.

很需要小心的是，原初核合成理论尚不能十分肯定地给出 Ω_B 的值(参看第六章). 这使得上面的推论只是启示性的，而不是结论性的. 如果我们相信原初核合成理论，它今天给出的预言是 Ω_B 等于 0.04 左右. 这进一步暗示，星系尺度上的暗物质就是以非重子为主的，暗晕中的气体主要不是电离的或非电离的原子气体. 类似地，星系团中会发出 X 射线的热气体也不是星系际气体的主要部分. 目前这些推测都已有了或强或弱的观测证据.

对非重子为主概念的另一方面支持来自结构形成的研究. 这问题属下一章的内容. 这里只提前指出，若宇宙中只有重子物质，

那么今天应来不及形成大量星系，从而是明显与事实冲突的. 非重子为主概念的出现，使结构形成研究者有柳暗花明的感觉. 因此结构形成问题是非重子为主概念第一个深深地扎根的领域.

六、热和冷的暗物质

当很多人意识到非重子物质才是宇宙介质中最主要的组分，那么它是什么的问题自然地得到了大家的重视. 从观测天文的眼光看，非重子物质的概念太超乎想像了，因此对只习惯于从观测获取知识的人感到很难接受. 在宇宙学的眼光看来，这却只是一个从早期演化中自然引伸出来的概念.

在宇宙的甚早期，自然界一切可能存在的粒子几乎都大量地存在过，不管这种粒子是稳定的或不稳定的. 随着宇宙温度的下降，当不稳定粒子失去了热碰撞产生的可能，它们将通过衰变而从宇宙中消失. 凡是稳定的(或寿命极长的)粒子，必定会在今天的宇宙中留下它的遗迹. 因此在重子物质外有其他非重子物质存在，完全是必然的事. 让我们分析不同粒子留下遗迹的过程，以了解它们对今天宇宙密度的贡献.

先讨论很轻的粒子，例如光子和中微子. 光子的退耦已专门讨论过. 现在着重讨论中微子. 中微子的退耦发生在 $T=1$ MeV 前后，而它的静质量远小于 1 MeV. 因此它是作为辐射粒子退耦的，退耦留下的粒子数密度很大. 由于它们在退耦时的热运动速度极接近光速(即很热)，所以被称为热暗物质. 若有其他退耦温度 $T_{dec} \gg M$ 的稳定粒子，它在退耦后也将成为今天的热暗物质.

然后看热暗物质对宇宙密度 Ω_0 的贡献. 不管是什么具体的热暗物质，其今天的温度总是绝对温度几度，其粒子数密度粗略讲来是 10^2 个/cm^3. 若它是像光子那样的零质量粒子，它对宇宙密度的贡献是 10^{-5}，即完全不重要的. 若它有一点静质量如中微子，那么它对 Ω_0 的贡献就具体地取决于其质量大小了. 记得临界密度的

大小是 $\rho_c = 10^4\ \mathrm{eV/cm^3}$. 要是其质量远超过 100 eV，这种热暗物质的贡献将达到 $\Omega_0 \gg 1$，这是宇宙学的事实所不容的. 具体看中微子. 按粒子物理的标准模型，中微子是零质量粒子. 近年已有较可靠的证据表明它其实有一点质量. 中微子有多大的质量?这不仅是粒子物理关心的问题，也是宇宙学关心的问题. 大致讲，每 1 eV 的质量将对 Ω_0 引起 0.01 的贡献. 若其质量远小于 1 eV，那么它作为暗物质是不重要的.

另一种情形是很重的稳定粒子，即 $M \gg T_{\mathrm{dec}}$. 这种粒子在退耦前已绝大部分成对地湮灭了，所以遗留下来的粒子数密度肯定很少. 考虑到它在退耦时的热运动速度已远小于光速，人们叫它冷暗物质. 我们类似地讨论它对宇宙密度的影响.

它的密度是其质量与粒子数密度的乘积. 这里特别需要注意的是，它残留的粒子数密度敏感地依赖 M 与 T_{dec} 之差. 湮灭阶段粒子数的减少是很迅速的. $M - T_{\mathrm{dec}}$ 越大，则退耦前的湮灭阶段越长，于是残留下的数密度就低了很多量级. 固定 T_{dec} 来考虑更清楚. 质量大了几倍，而残留的粒子数会少几个量级. 于是它对 Ω_0 的贡献反而小了很多. 这样导致了一个佯谬般的结论：冷暗物质粒子越重，它对今天宇宙密度越次要. 按 $T_{\mathrm{dec}} = 1$ MeV 来分析，若冷暗物质粒子的质量显著低于 10 GeV，将导致 $\Omega_0 \gg 1$. 与上面讲的一样，这是宇宙学事实不允许的. 反之，若其质量远高于 10 GeV，那么这粒子作为冷暗物质是很次要的.

这里出现的戏剧性局面是，按已发现的 800 多种粒子[①]，没有一种能作为冷暗物质粒子出现；可是在高能物理理论中，候选者是存在的. 这里最被人们注意的是超对称理论预言的普通粒子的超对称对应物. 这类粒子中最轻的中性粒子应是稳定的，它今天应作

---

① 其中绝大部分都是寿命非常短的粒子. 已知的稳定或长寿命粒子只有质子、中子、电子、中微子和光子.

为冷暗物质存在，它的名称叫 neutralinos①. 人们希望它的质量比 10 GeV 重，但是重得不太多，那么它将是极重要的冷暗物质粒子. 另一种被注意的是轴子(axions). 它是强 CP 理论中预言的粒子.

现在人们热衷于研究的问题是如何发现它们，如果它们是现实的冷暗物质粒子，应该会引伸出什么可观测的现象，等等. 目前离问题的澄清还差得很远，但各种手段的研究在认真进行中.

有可能会出现中间状况，即某种稳定粒子的质量与其退耦温度接近相等. 这被叫作温的暗物质. 这种情况也是有人研究的. 我们在此不再讨论.

七、问题的小结

从前面的讨论看到，暗物质是与宇宙学中许多问题有关的枢纽性问题. 此外这问题也有着很丰富的物理内容，并对粒子物理有所启示. 因此它吸引了各方面人的投入，既包括宇宙学和粒子物理的理论研究者，也包括观测和实验的研究者. 经多方面的多年研究，人们对问题的答案形成了一些猜测性的看法，可是宇宙学需要的是由事实作出的回答.

暗物质研究想澄清的两个最基本的问题是：(一) 宇宙总密度到底多大？(二) 宇宙物质的最主要组分是哪几种？直到最近两三年，人们才能够对此给出一份初步有事实根据的答卷. 我们把它归结为如下几点.

(一) 宇宙的总密度很接近于 1. 它是符合暴胀理论的预言的. 它究竟比 1 略大还是略小，这关系到宇宙的有限或无限，因而受到某些哲学家的关心. 其实这问题对于宇宙学或物理学都没有任何重要性. 由于实际测量范围内密度必有涨落，以及测量必有误差，

---

① 直译为中性的费米子. 这明显不是确切的名称. 其实指的是普通中性规范粒子和中性 Higgs 粒子(都是玻色子)的超对称对应物.

因此在相当长的时间里，我们难望能有清楚的答案.

（二）真空是宇宙的最主要的组分. 它的密度约占总密度的65%左右. 它决定了今天的宇宙已转入加速的膨胀阶段. 在宇宙年龄为 0.1 t_0(即 15 亿年)之前，这部分能量对宇宙的膨胀几乎没有影响. 它影响的只是星系开始形成以来的时期.

（三）宇宙中最主要的实物是非重子物质. 它约占总密度的30%左右. 星系形成的研究表明(见下一章)，冷暗物质应是主要的. 热暗物质只能占相当小的比例，或者完全不重要. 至于冷暗物质是什么，问题还远不清楚.

（四）普通由原子组成的物质比冷暗物质少很多. 它应只占总密度的 5%或更少. 重子物质在宇宙中的重要性在于没有它就没有人类，没有它就不会有我们赖以生存的恒星和行星①. 在更大的尺度上，看不到它对宇宙的行为还有重要的影响. 当然，没有人类也就不会有宇宙学了.

这就是人们今天的认识. 当然我们最好仍保持开放的头脑，把这些仅看成能得到 70 分以上的答卷. 实际与上述回答有差别是可能的. 无论如何，经过近 30 年的努力和期盼，我们离最终的答案已经不很远了.

---

① 由于非重子物质没有通过辐射来耗散能量的机制，它不能形成如恒星这样致密的天体. 冷暗物质能有星系级的结团，热暗物质可能只有超星系团级的结团，所以要是没有重子物质，宇宙的结构会简单得多.

第十一章　结构的形成

今天宇宙物质已有层次性的结团.这是最基本的事实.大爆炸理论认为早期宇宙十分均匀,观测也已证实了它.那么,如何从高度均匀的气体中演化出层次性的结构,这无疑必须回答.从理论上讲,要回答这问题需事先知道宇宙的密度、介质的组分,以及密度起伏的初条件.长期以来,由于这三个前提都不具备,人们一直只能作试探.这样倒也澄清了不少疑点,获得了不少经验.最近几年对此问题的研究有很大进展,前两个前提已大体清楚,后一要求即将有新的观测提供信息.乐观地看,人们艰辛的努力快要结出果实了.

一、自引力的不稳定性

牛顿就意识到,万有引力的存在会造成物质的聚积.他在致Bentley的信中曾说,如果太阳、行星和一切物质曾均匀地弥散于宇宙空间,由于物质间固有的引力,物质将局域地集积.在无限的空间中,它们将形成无数个大的物质团(大意).在他看来,满天星斗就是这样造成的.牛顿在这些话中表述的,正是我们现在要讨论的概念:引力会造成均匀介质的不稳定,从而发生物质的局域结团.

这方面的具体理论是在20世纪初,由Jeans首创的.Jeans考

虑一大片静止而均匀的流体介质，其密度记作 ρ_0. 当有某种机制在介质中造成了微小的不均匀，其密度变为

$$\rho(\boldsymbol{r},t=0)=\rho_0+\rho_1(\boldsymbol{r},t=0), \tag{11.1}$$

他要用流体动力学规律回答，介质各部分间的自引力将怎么使扰动量 ρ_1 演变.

由流体动力学可导出，微小密度起伏 ρ_1 的演变应满足方程

$$\frac{\partial^2\rho_1}{\partial t^2}-v_s^2\nabla^2\rho_1=4\pi G\rho_0\rho_1, \tag{11.2}$$

其中 v_s 是均匀介质中的声速. 式中出现含引力常数 G 的项，反映自引力的效果. 若去掉这项，方程将还原为标准的声波方程. 这表明若自引力不重要，扰动将以声波的方式传播. 现在感兴趣的正是自引力的影响.

按照傅里叶分析的数学理论，任意的密度扰动都能展开为正余弦型扰动的叠加. 为方便而采用复数写法，即有

$$\rho_1(r,t)=\frac{1}{(2\pi)^{3/2}}\int\rho_k(t)e^{ik\cdot r}d^3\boldsymbol{k}, \tag{11.3}$$

物理的扰动量理解为其实部. 这样我们可以只讨论一种波矢 $\boldsymbol{k}$ 确定的余弦型初始扰动的演化，而并不失去一般性.

对这类型的扰动，其演化由扰动幅度 $\rho_k(t)$ 描述. 它满足的方程是(来自(11.2))

$$\ddot{\rho}_k(t)=-(k^2v_s^2-4\pi G\rho_0)\rho_k(t). \tag{11.4}$$

上加的附点代表对时间的微商. 易于把(11.4)解出，

$$\rho_k(t)=\rho_k(t=0)e^{i\omega t}, \tag{11.5}$$

其中 ω 是圆频率. 方程要求它与波矢 k 有关系

$$\omega^2=k^2v_s^2-4\pi G\rho_0.$$

现在可清楚地讨论自引力的影响了.

我们把密度超过 ρ_0 的地方叫正密度扰动区. 这区域的大小大体相当于半波长 $\lambda/2$(图 11.1)，即 π/k. 当正密度扰动区很小(即波矢 k 很大)，(11.6)式右边第二项小于第一项，因而 ω^2 大于零.

这表明自引力起次要作用时，扰动仍以声波的形式传播. 扰动的幅度保持不变. 反之若正密度扰动区很大，第二项将大于第一项，从而使 ω^2 小于零了. 这是自引力起主要作用的情形. 从(11.5)式看出，ω 是虚数说明扰动的幅度将按指数律迅速增大，这就是自引力的不稳定性. 它的含义是：均匀介质中一旦出现小的密度起伏，自引力将使起伏的幅度放大、引伸下去，最后将局部地形成密度远超过 ρ_0 的物质团块. 这就是宇宙中物质结团的基本机理.

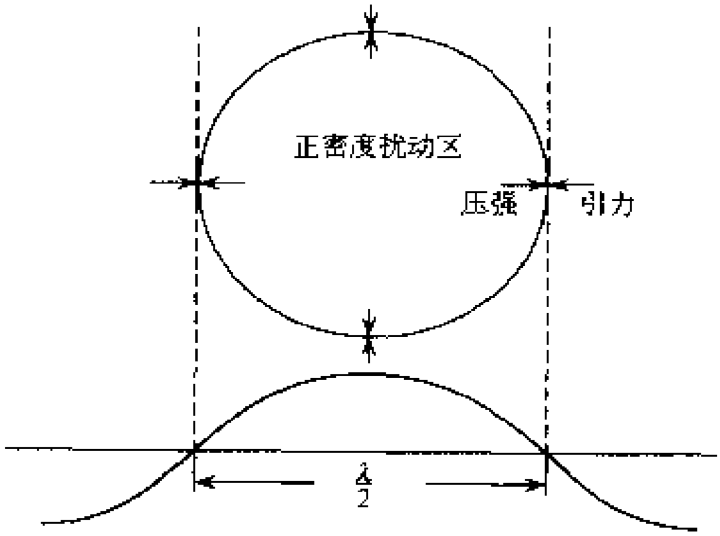


图 11.1　正密度扰动区

让我们更物理地理解结团的起因. 参照图 11.1，正密度扰动区内压强也随密度而增大了，它倾向于把区内的物质向外推. 这是形成声波的机理. 自引力的存在造成了一种相抗衡的因素. 区域内密度增大对周围物质产生了更大的引力，从而有使这区域内密度更增大的倾向. 于是，这两种相反的倾向出现了竞争. 压强把内部物质向外驱赶的时间尺度是 $\tau_{压} \sim \lambda/v_s$，即用声速把物质推出这扰动区的时间. 自引力使区域内密度变得远超过 ρ_0 的时间尺度是 $\tau_{引} \sim (G\rho_0)^{-1/2}$. 对照(11.6)式看，两种机制的竞争就是 $\tau_{压}$ 和 $\tau_{引}$ 比大小. 自引力为主时 $\tau_{引}$ 很小，于是在压强还来不及起作用时，区域内的密度已显著地增大了.

Jeans 理论不仅说明自引力会导致不稳定性，而且指出，不稳定性的出现是有条件的. 这条件就是(11.6)式右边要小于零. 相应的临界波长被称为 Jeans 尺度，记 λ_J，

$$\lambda_J = \left(\frac{\pi v_s^2}{G\rho_0}\right)^{1/2}. \tag{11.7}$$

这尺度的大小取决于介质的密度和温度(决定声速). 人们也常用 Jeans 尺度内的物质总量 M_J(叫 Jeans 质量)代替 λ_J，作为临界的标志.

$$M_J = \frac{4\pi\rho_0}{3}\left(\frac{\lambda_J}{2}\right)^3. \tag{11.8}$$

这样可把 Jeans 理论的结论表述如下：当有正密度扰动区，其内的质量超过 Jeans 质量，则这扰动是自引力不稳定的. 或换句话说：Jeans 质量是介质中能因小扰动而造成结团的最小质量尺度.

二、膨胀宇宙中的小扰动

Jeans 理论是静止气体云中形成恒星的理论，它不直接适用于宇宙. 这里有两点欠缺. 一是没有考虑到介质在膨胀，二是对超过视界的扰动区须用广义相对论代替牛顿理论. 这两个问题已在 1946 年由 Lifshitz 完善地解决了. 我们只扼要地讨论它与 Jeans 理论的差别.

为简单起见，下面仍讨论亚视界扰动区，以便继续使用牛顿的理论①框架. 我们的主要兴趣在于弄清膨胀对扰动演化的影响.

在膨胀介质中，均匀的背景密度 ρ_0 也随时间而变化. 这变化由 Friedmann 方程决定，因此是在研究扰动演化前已经清楚的. 这样人们引入相对密度

$$\delta(\boldsymbol{r},t) \equiv \rho_1(\boldsymbol{r},t)/\rho_0(t) \tag{11.9}$$

---

① 当区域尺度远小于当时的视界，牛顿理论是广义相对论的好的近似.

来描写扰动.它的意义是扰动引起的密度反差.

现对密度反差 $\delta(\boldsymbol{r},t)$ 做傅里叶分解,得到分波振幅 $\delta_k(t)$. 用几乎一样的方法能从流体动力学导出 $\delta_k(t)$ 满足的方程.它是

$$\ddot{\delta}_k + 2\frac{\dot{R}}{R}\dot{\delta}_k + \left(\frac{k^2 v_s^2}{R^2} - 4\pi G\rho_0(t)\right)\delta_k = 0, \tag{11.10}$$

其中的 R 是服从 Friedmann 方程的宇宙尺度因子,$\rho_0(t)$ 是变化的背景密度.这里的 k(或相应的 λ)是固定的随动[①]波矢(或波长).意思是所考虑的扰动区内的物质总量固定,其物理大小 $R(t)\lambda$ 随膨胀而增大.把这方程与 Jeans 理论中相应的方程比较,主要的差别是多了一个正比于 $\dot{\delta}_k(t)$ 的项.这是振动方程中的阻尼项.由此看来,宇宙膨胀对扰动演化的影响是产生了一种阻尼.

仍用方程(11.10)的第三项为零来定义 Jeans 尺度 λ_J,

$$\lambda_J = \frac{2\pi}{k_J} = \frac{1}{R(t)}\left(\frac{\pi v_s^2}{G\rho_0}\right)^{1/2}, \tag{11.11}$$

它从形式到内容都与 Jeans 理论中的定义是一样的.当扰动的范围很小($\lambda<\lambda_J$),方程(11.10)中第三项的系数为正.这时方程有阻尼振动解.因此原来会以声波传播的扰动,现在它仍将以声波在传播,但是将在传播中逐渐阻尼掉.我们特别感兴趣的是原来会出现自引力不稳($\lambda>\lambda_J$)的情况.现在需要弄清膨胀阻尼的表现是什么,以及它有多大的影响.

为从方程(11.10)中解出扰动的演化,需要先把均匀背景的行为代进去.在辐射为主的早期和在实物为主的后期,宇宙的膨胀行为不同,因此扰动的演化也会不同.它们需要分别讨论.

把实物为主时期的宇宙行为代进去.因为我们只想讨论自引力为主的情形,所以把压强项 $k^2v_s^2/R^2$ 忽略,扰动方程(11.10)具体地化成

---

① 这相当于对随动坐标做傅里叶分解.

$$\ddot{\delta}_k + \frac{4}{3t}\dot{\delta}_k - \frac{2}{3t^2}\delta_k = 0. \tag{11.12}$$

这二阶微分方程的两个独立解为

$$\delta_k(t) = \text{const.}\, t^{2/3} \tag{11.13}$$

和

$$\delta_k(t) = \text{const.}\, t^{-1}. \tag{11.14}$$

这表明自引力起主要作用时依然有不稳定性出现，但是扰动的增长显著减慢了．它只与时间 t 的 2/3 次方幂成正比．

辐射为主阶段须看成双组分气体．光子是主要组分，但是我们感兴趣的是实物组分上的扰动．按这情况，实物密度反差满足的方程可同样地导出．它的形式是

$$\ddot{\delta}_k + \frac{1}{t}\dot{\delta}_k = 0. \tag{11.15}$$

这方程的两个独立解是

$$\delta_k(t) = \text{const.} \tag{11.16}$$

和

$$\delta_k(t) = \text{const.}\ \ln t. \tag{11.17}$$

这结果表明自引力不稳定性依然出现，但是这么慢的增长实际上等于没有增长．因此早期存在着的小扰动须等待到实物为主开始后，才有有效的放大．

从这些计算结果看到，宇宙膨胀对自引力不稳定性的影响是使小扰动增长的速率大大地减慢了．在辐射为主阶段扰动实际上没有增长．实物为主阶段①约开始于 $t=10^5$ 年，此后时间每过三个量级，小扰动的幅度才增大两个量级．这样，物质的结团将发生在很晚近的阶段．

值得注意一下如果宇宙是低密（Ω_0 比 1 小得多）的情形，把后期实物为主阶段的 Friedmann 方程写成

---

① 实物为主开始时刻 $t_{eq}=1.3\times10^3(\Omega_0h^2)^{-2}$年，注意它与 Ω_0 平方成反比．

$$\frac{1}{H_0^2}\left(\frac{\dot{R}}{R}\right)^2=\frac{\Omega_0 R_0^3}{R^3}-\frac{k}{H_0^2 R^2}.\qquad(11.18)$$

若其右边的后项比前项小，这叫曲率为主阶段. 把上式用于今天，其左边是 1，因此

$$\Omega_0-\frac{k}{H_0^2 R_0^2}=1.\qquad(11.19)$$

只要今天有 $\Omega_0>0.5$，曲率为主将发生在将来，从而不必考虑. 要是今天的密度仅为例如 0.1 到 0.2，那么在今天前的一个晚近阶段是曲率为主的. 我们不再写公式而指出相应情况下的结果：曲率为主阶段小扰动也是不会增大的. 注意一旦进入曲率为主，曲率项将越来越大于密度项. 因此，某尺度上的小扰动在进入曲率为主时若还没有足够地被放大，那么它将不再能增大，也不会发展成物质结团了.

三、重子宇宙模型的困难

直到 70 年代末，人们自然地认为实物就是由重子组成的普通物质. 原初核合成的研究成果又使人们意识到，宇宙密度（即重子物质密度）不能太大. 下面即将论证，在这样的前提下，星系竟然是至今来不及形成的. 这是用重子宇宙模型研究结构形成的重大困难.

让我们具体地讨论 $\Omega_0=0.1$ 的重子宇宙模型[①]，以看清这困难的由来.

在这样的模型中，实物为主阶段开始于 $R=10^{-3}R_0$ 时. 这时宇宙年龄接近 10^6 年，温度为 2700 K. 光子与实物的退耦也发生在差不多同一时刻. 容易由定义算出 Jeans 的临界质量随温度的变化，结果如图 11.2 所示. 当实物与光子尚有耦合，实物感受到的

① 当然须认为宇宙中还有光子，否则就不会有辐射为主的早期了.

压强主要来自光子，因此介质中的声速 v_s 很大. 它是光速的 1/3. 在光子退耦后，实物介质中的声速远低于光速. 这声速的骤然降低是造成 Jeans 质量有一向下跳变的主要原因.

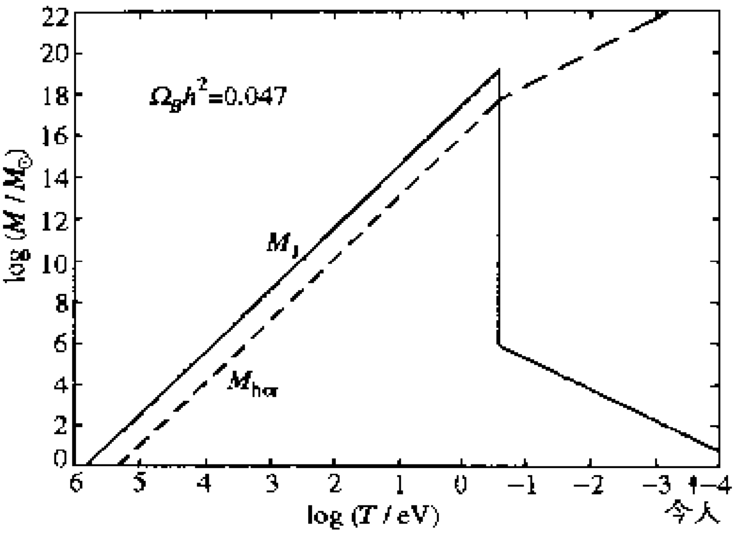


图 11.2　重子物质的 Jeans 质量

我们记得，扰动在实物为主期才增长，而且扰动区内的质量又须大于 Jeans 质量. 此外上述讨论还有一个前提：扰动区的大小须小于视界(视界内的质量也已画在图中，记作 M_{hor}). 这样可从图上看出，最先增长的扰动尺度①是 10^6 到 $10^{17}\ M_\odot$. 这里重要的问题是：这些小扰动到什么时候能发展到局域的结团.

结团过程可以粗分为两段. 前一段是从小扰动($\delta_k \ll 1$)发展成"大"扰动($\delta_k \sim 1$). 这段可用上面的方法处理. 到扰动区内的密度已显著高于周围的环境，环境的影响已变得次要了. 它会由内部的引力而向自己的中心坍缩，直至靠内部压强而重建力学平衡. 这样它才形成一个密度远高于环境的系统.

对低密度的重子宇宙模型，从实物为主开始时刻($t \approx 10^6$ 年)

① 考虑到 Silk 阻尼，最先增长的尺度是从 10^{14} 到 $10^{17}\ M_\odot$. 这问题放在后面讨论.

起到曲率为主时刻($t \leqslant 1\times10^{10}$年)止,小扰动只能增大不到3个量级,此后它就不再增长了.从背景辐射上测到的温度起伏推断,当时的密度反差仅为$\delta_k<10^{-3}$.这样看来,允许扰动增长的时间太短了.今天($t \geqslant 1\times10^{10}$年)小扰动还没有增长到$\delta_k \sim 1$,从而更来不及有星系的形成.这就是星系形成的时间困难.

这里也讨论一下高密的($\Omega_0 \sim 1$)重子宇宙模型.在这样的模型中,实物为主开始得要早两个量级,即接近$t=10^4$年.此外晚近又不会出现曲率为主.这样允许扰动增长的阶段延长了很多.但是星系形成的时间困难却只得到不大的缓解.

虽然实物为主提前了很多,但是光子的退耦时刻①并没有提前多少.它约发生在$t=2\times10^5$年时.上面已提到,在光子退耦前的Jeans质量很大(参考图11.2).所以虽已进入实物为主,各种天文尺度尚小于Jeans的临界尺度,扰动的增长还须等待到光子退耦之后.这样要在今天之前有星系的形成,仍然非常勉强.此外重子密度达到$\Omega_0 \sim 1$的前提与原初核合成理论有冲突.看来,这模型不像是现实的写照.

顺便看高密度的非重子宇宙模型.非重子(例如中微子)与光子没有耦合,所以它的Jeans尺度下降较早.在进入实物为主后,星系和星系团尺度的扰动区很快超过了非重子物质的临界尺度,从而开始向"大"扰动发展.因此在这样的模型中,时间困难得到了很大的缓解.这就是非重子为主的概念有很大吸引力的首要原因.当然我们感兴趣的是重子物质能及时地结团.

非重子物质中的小扰动从实物为主起即开始增长,而重子物质中的小扰动还不能增长.图11.3画出了这情景.到光子退耦发生,重子扰动才开始增大,但是这时密度反差已增大的非重子物质对它产生了附加的引力,使重子扰动的增大加快了.逐渐地,两者的演化会同步起来.于是重子物质也能跟非重子物质一起及时地

① 原子中性化发生在$t_{rec}=1.8\times10^5(\Omega_0 h^2)^{-1/2}$年.它与$\Omega_0$开根成反比.

结团. 这样，星系形成的时间困难才是消除了.

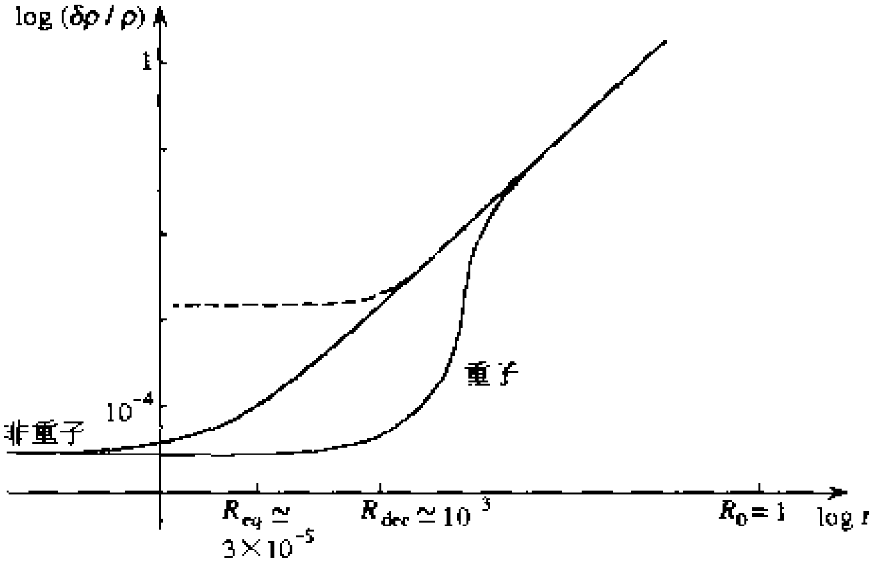


图 11.3　混合气体中扰动的演化

四、结构形成作为初值问题

从理论的角度讲，结构形成是扰动演化的初值问题. 问题的一方面需要先弄清原始扰动的产生机制是什么，以及所产生的扰动幅度 δ_k 有多大. 这就是以后演化的初条件. 另一方面要弄清扰动演化的动力学规律(即微分方程)是什么. 掌握了这两方面，剩下的就是用数学方法求解初值问题了. 把微分方程积到今天，就应该能得到今天物质结团的状况. 但是一直到 70 年代末，用这办法系统地研究结构形成的条件完全不具备.

首先遇到的问题是第八章中讲过的视界疑难. 追溯到早期，星系尺度的扰动区是超视界的. 这使得人们原则上不可能找到产生原始扰动的物理机制. 这也就是说，星系的起源问题物理上不能研究，扰动演化的初条件是理论上得不到的. 为避开这原则性的困难，人们把任何尺度的扰动进入视界的时刻当作实际研究的初时

刻.此后这些扰动永远将是亚视界的,它的动力学演化才能够用物理方法研究.

图 11.4 中画出视界内的质量随时间的变化.实物为主的开始时刻记作 t_{eq}. 这时刻视界内的质量大体与星系团相当.这就是说,星系团尺度的扰动是在实物为主开始时进视界的.星系尺度要小 1 个量级以上.它是在辐射为主的后期进视界的.同理,超团或更大尺度的扰动是在实物为主后才进视界.如果我们能知道每一尺度(用 k 或 λ 标记)的扰动进视界时的幅度 δ_k,它就可当作扰动演化的初条件.这样的初条件叫进视界扰动谱.

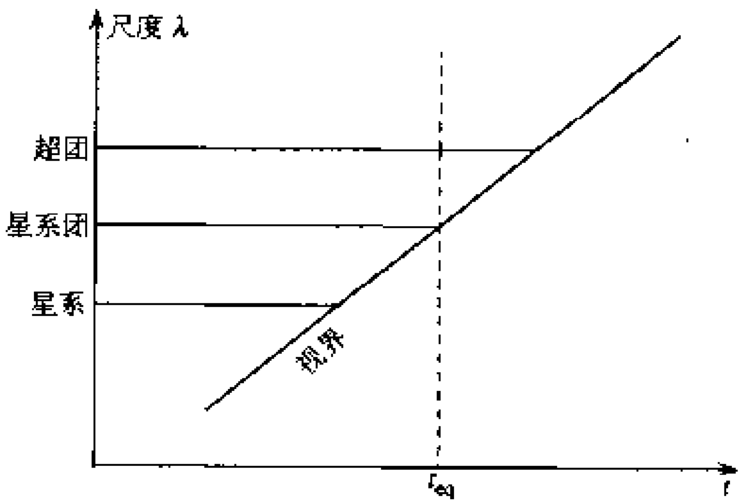


图 11.4　天文尺度的进视界时间

70 年代初,Harrison 和 Zeldovich 分别通过唯象分析,作出了同样的推测:一切尺度的扰动在进视界时有同样的扰动幅度,即

$$\delta_k(\text{进视界}) = \text{const.}. \tag{11.20}$$

文献上常叫它为尺度无关谱,或 Harrison-Zeldovich 谱.这幅度的大小约是 10^{-5}.注意这进视界扰动谱仅是有一定依据的猜测.

研究扰动的演化靠的是动力学方程.要使这方程具体化,须事先知道宇宙模型的总密度,以及宇宙介质中各主要组分所占的比

例.在认识到非重子物质的重要作用前,自然的选择当然是重子加光子,以及今天以重子为主的模型.在这样的模型呈现出时间困难时,人们感到困惑莫解.想不到错误可能出在哪儿.因此,当有迹象表明中微子可能是宇宙物质的主角时,结构形成的研究者有了豁然开朗的感觉.那时已到80年代初期了.

80年代初期同时出现的另一进展是暴胀理论的提出.它指出在暴胀前,任何天文尺度的扰动都是亚视界的,因而原初扰动应产生在暴胀之前.有了非重子为主以及甚早期暴胀这两个新概念,结构形成的研究出现了全新的局面.从这时起,把结构形成当作扰动演化的初值问题来研究已有了可能.不过需要强调,暴胀是否确实发生过是一个疑问,非重子暗物质是否是主要组分也是疑问.因此我们只能说,结构形成问题的研究由此才开始走上了正路.至于要问题得到解决,人们还有漫长的路要走.人们不能期望一下子从暴胀末产生的原始小扰动出发,就能算出今天宇宙的全部面貌.这仅是最终目标.实际的研究是三个方面独立地、相互参照地发展着.

从背景辐射上测定当时的扰动幅度是一个方面.背景辐射上反映的是t_{dec}时刻的密度起伏.这时刻比实物为主的开始时刻t_{eq}晚一些.若能测定$\delta_k(t_{dec})$,那么计算以后扰动的演化就有了可靠的初条件.另一方面,理论上能把$\delta_k(t_{dec})$和δ_k(进视界)联系起来,从而可以知道Harrison和Zeldovich关于尺度无关谱的猜想是否近似地对.这方面的实测研究是结构形成问题中最基本的部分.

在掌握可靠的初条件前,用猜测性的初条件研究其演化后果是另一个很大的领域.

这种研究先不问扰动的来源,只管它们进视界后的演化.目标是在猜想性的前提下得到符合现实的宇宙面貌.从80年代初期起,人们的兴趣是弄清,若宇宙以热暗物质为主,或以冷暗物质为主,能否正确地演化出今天的宇宙.我们将在后面两节中讨论这方面研究的收获和疑问.

天体物理学家着重的是扰动演化的后一半,即进视界后的一

半. 粒子理论学家则更感兴趣于扰动进视界前的一半. 这问题包括两方面. 首先是研究暴胀如何产生原始扰动和算出相应的扰动谱. 这是结构形成真正的(物理的)初条件. 注意到扰动一经产生,这尺度立刻越出了视界①. 于是进一步需要研究,超视界的扰动如何按广义相对论的规律演化. 研究的目标是: 希望在各扰动尺度再进入视界时,它的幅度 δ_k(进视界)能符合后面演化的需要,或符合实际测到的结果. 这方面的研究状况我们不准备深入讨论了.

五、热暗物质为主的模型

现实存在的热暗物质就是中微子. 它是否是主要组分,取决于它的静质量. 从结构形成问题讲,任何一种热暗物质的行为是差不多的. 所以,我们抽象地讨论热暗物质为主的模型. 它假定热暗物质占例如 90%,重子物质相应地占 10%,总密度为 $\Omega_0=1$,当然光子作为次要组分也存在.

再假定各种尺度的扰动在重进入视界时有相同的幅度,即采用 Harrison-Zeldovich 谱作初条件,那么此后扰动的演化就能计算了. 由于各种尺度的扰动区进视界的时刻不同,这不是一个等时的初条件. 可以用理论方法把它们折合到同一时刻——实物为主的开始时刻 t_{eq},得到的等时初始扰动谱如图 11.5 中的 A 线所示. 但是这里还有一个物理机理没有考虑进去.

热暗物质作为已退耦的无碰撞气体,粒子的自由流动会削弱不同地方的密度差别. 这在物理上叫无碰撞阻尼,或叫朗道阻尼. 在辐射为主时期,热暗物质粒子的热运动速度几乎就是光速. 因此在退耦以来长时间的自由流动中,星系和星系团尺度的密度起伏几乎完全被抹平了. 到实物为主开始时,剩下的只有超团和超团以

---

① 从图 8.3 看到,暴胀中星系等一切天文尺度都将随观测宇宙越出视界. 而后因视界增长得快,于是在宇宙转入实物为主的前后,各种尺度又"重新进入"了视界,如图 11.4 所画. 这里把演化分前后两半,就是以重进视界时刻为界而分的.

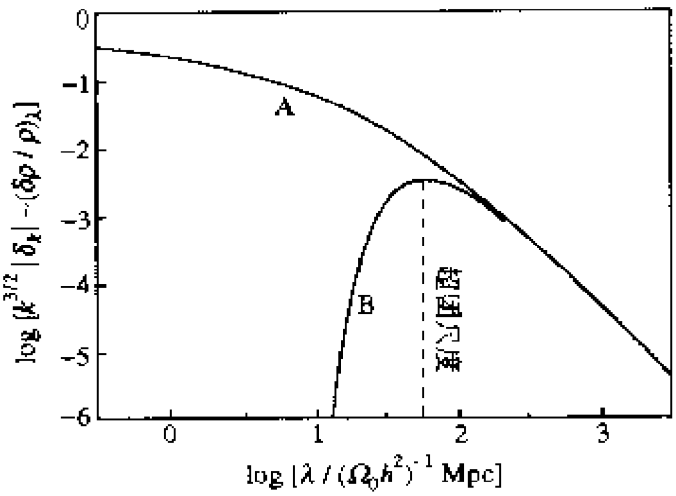


图 11.5　等时的初始扰动谱

上尺度的扰动，如图中的 B 线所示. 这才应是热暗物质中扰动演化的初条件.

按这样的初条件，首先出现的将是超团质量的物质团块. Zeldovich 早年论证过，当这么大的气体云发生自引力坍缩时，有一维会优先缩小. 所以超团不呈球状，而呈"薄饼"状. 在热暗物质为主的模型中，这坍缩是在混合气体中发生的. 由于重子物质能通过辐射，部分地把坍缩增大的热能耗散掉，于是它能收缩到更薄. 热暗物质没有耗散能量的机制，它停止收缩时的致密程度低一些. 于是所形成的薄饼呈三文治(sandvich)状，重子物质夹在中间.

重子物质在这坍缩中会因各种不均衡而发生碎裂，以结成尺度更小，密度更大的团块. 这就是将来的星系和星系团. 热暗物质在坍缩后已高度简并①，不可能再碎裂成密度更大的小团块了. 因此最终形成的局面是，重子物质结成了由星系和星系团组成的超团，而热暗物质将以较小的密度分布在几十 Mpc 的范围内，成为大量星系共有的暗晕. 结团过程是自大而小地进行的. 这是热暗物质模型中结构形成过程的重要特征.

---

① 受泡里不相容原理的约束，中微子都占据了尽量低的能态叫高度简并. 这样再增大密度需要外界提供能量，以使粒子去占据更高的能态.

这模型中占最主要比例的热暗物质以很弥漫的超团暗晕存在. 这样就很自然地说明了,在总密度为 $\Omega_0=1$ 的宇宙中作局域测量,只会测到一个很低的密度,例如 $\Omega_0=0.2$. 这是第一个成功之处. 对结构形成过程作计算机模拟,至演化后期看到出现了大的空洞,以及星系呈纤维状分布在洞壁上. 大尺度(尺度>10 Mpc)的天体分布与从巡天观测得到的分布图景很相似. 这又是热暗物质模型的成功之处. 但是在中、小尺度上的情况很不好.

在第一章第三节中讨论过,尺度为 λ 区域内的物质起伏可用 $\langle\delta M/M\rangle^2$ 的全宇宙平均值描述. 观测表明 $\langle\delta M/M\rangle^2=1$ 的尺度约为 8 Mpc, 尺度越大它的值越小. 这表明全宇宙平均地讲,大尺度的扰动今天尚远没有发展成熟,而在星系和星系团尺度上,扰动的发展要成熟得多. 热暗物质模型中首先从大尺度(~40 Mpc)开始结团,中小尺度结团发生在后. 这样,模型在中小尺度上的现象很可能会有不够成熟的毛病.

数值模拟表明,超团型团块的形成已较晚,约在红移 $z\leqslant 3$. 星系作为它进一步碎裂的后继产物,约形成于 $z\leqslant 1$. 实际上红移超过 2 到 3 的星系已发现了很多. 若要在这模型中让星系形成于 $z=2$ 之前①,将会使大尺度结构过分成熟. 用模拟研究星系与星系的空间关联或其他方面时也出现同样的毛病. 总之由于热暗物质模型中最早启动的扰动尺度太大,中小尺度上的这些毛病多少是在意料之中的.

六、冷暗物质为主的模型

现在看在冷暗物质为主的模型中,所形成的物质结团能否与现实符合得更好. 我们仍讨论 $\Omega_0=1$ 的高密模型,让冷暗物质约占

---

① 图 11.5 示出的仅是不同尺度上原初扰动的相对大小,即扰动谱的形状. 因尚没有实际测定,扰动的绝对大小依然可以调节.

到 90%. 在指定了模型宇宙的密度和主要组分后,尚需要的是指定扰动的初条件.

依然采用 Harrison-Zeldovich 的尺度无关谱,并把它折合到 t_{eq}时刻. 等时的初始扰动谱仍是上图中 A 线的形状. 这里开始出现了与热暗物质模型的重要差别. 冷暗物质粒子退耦时的热运动速度已远小于光速,所以退耦后的自由流动距离要小得多. 这意味着被无碰撞阻尼所抹平的扰动尺度很小. 算出的结果表明,它比星系尺度还小很多,因此对实际感兴趣的结团没有影响. 这差别决定了其结团过程与热暗物质模型中的过程将很不一样.

按这样的初始扰动谱(图上的 A 线),越小的尺度上扰动幅度越大,因此亚星系尺度的结团将率先发生. 亚星系团块形成后将再聚合成星系,一群星系再聚合成星系团,直至超团. 整个结团过程按先小后大的顺序进行. 这叫逐级结团图景. 它与热暗物质模型中的图景是相反的.

80 年代中期出现了一系列很漂亮的数值模拟工作,分别研究了小、中和大尺度上的结团. 小尺度指几个 Mpc 的正密度扰动区,即能形成一个星系的范围. 模拟能得出星系形成后的内部物质分布,以及星系的转动曲线等. 所得结果与事实符合很好. 转动曲线在外部确平坦地延伸,如同上章指出的那样. 中尺度指 10 来个 Mpc 的范围. 其中会形成好多个星系. 模拟能研究星系的数密度、星系在空间分布上的关联等. 星系-星系关联有很可靠的观测资料,模拟与事实也符合得很好. 可是在大尺度的模拟上,冷暗物质模型出现了明显的缺点.

大尺度指上百个 Mpc 的范围. 观测看到有很大的大空洞,例如牧夫座空洞的直径达到几十 Mpc. 模拟结果中却看不到. 甚至大尺度上很特征性的的泡抹状结构在模拟中也不明显. 研究星系团-星系团的关联,模拟出现的振幅至多只有实际的 1/3. 这些差别都表明冷暗物质模型中的大尺度行为较之实际尚不成熟.

此外值得注意,上面讲到的星系尺度结团,指的是冷暗物质与

重子物质的混合气体的结团，而还不是实际星系的形成. 在结团中气体温度要升高. 重子组分能通过辐射把部分热能抛出，从而它将与冷暗物质分离，集中到团块的中心. 这就是将来的恒星形成区. 冷暗物质无法进一步收缩，它将最终表现为恒星区外部的暗晕. 这才是在冷暗物质模型中星系形成的图象. 按这图象看，若宇宙总密度为 $\Omega_0=1$，在星系或星系团尺度上对密度作局域测量时，似不应只测到 $\Omega_0\sim0.2$. 当时许多研究者也注意到了这个问题，并提出了可能的解释. 在十来年后的今天看，这或许应看成是我们的高密宇宙以真空为主的迹象.

总之，冷暗物质模型中的结构形成在中小尺度上与实际的符合程度给人以深刻的印象，这是它很大的成功. 但是它在大尺度上演化的不成熟，也使人感到这模型还不是真实宇宙的写照.

七、近年的进展和展望

上面讨论到的是至 90 年代初的进展. 当时人们已意识到，非重子物质为主的模型能使我们粗略地理解有结构的真实宇宙的由来，但是不管以热的或冷的物质为主，模型也还都难以与现实相符. 回想 70 年代时人们连一个关于星系形成的试探性理论都没有，80 年代的进步是巨大的和实质性的. 但是这进步的背后尚有着很大的漏洞：理论的所有基本前提都没有证据.

最基本的前提是结构来自小扰动. 可是到 90 年代初，人们还没有从微波背景上找到小扰动的存在证据. 当时由找不到小扰动而给出的上限是 $\delta T/T<10^{-5}$. 若探测精度再提高一个量级而依然得到零结果，这漏洞将是致命的. 1992 年春天，COBE 卫星宣布找到了小扰动的信号. 这无疑是关键性的支持. COBE 测量的是张角为 10° 范围的温度起伏. 它相应于超过超团尺度的小扰动. 因此这发现的意义虽然十分巨大，但是把它作为研究扰动演化的初条件还是远不够的.

接着的进展是开始出现了三重主要组分的模型. 除重子物质不可缺少外,人们假定非重子物质中既有热暗物质,也有冷暗物质. 目的是想弄清,这样能否既保持冷暗物质模型在中小尺度上的成功,又在大尺度上具有热暗物质模型的优点. 人们按已有的经验,常假定热暗物质占最主要的比例(例如 70%),冷暗物质其次(如 25%),余下的才是重子物质. 总密度仍保持为 $\Omega_0=1$. 用模拟做演化的结果虽然不出乎意外,也不尽如人意. 于是有人做了用真空代替热暗物质作为主要组分的模型. 真空作为弥漫分布的暗物质,对结构形成的影响与热暗物质有类似之处. 当时真空能是否存在尚不清楚,有趣的是模拟发现,令真空代替热暗物质的模型确实比其他模型更好. 但是冷暗物质同时占较大比例是不可缺的. 这是在 90 年代中期又前进了一步的标志.

理论模型与现实又靠近了一步,模型的基本前提是否可靠的问题显得更突出了. 需要事实澄清的前提有几个: 1. 宇宙密度是 $\Omega_0=1$ 吗? 2. 宇宙确以非重子为主吗? 3. 真空确有能量吗? 在以此为前提的理论研究已取得了不少成功时,前提本身的可靠性竟然还完全没有证据!

这方面的进展在前面都已讲到过了. 在最近的两三年中,上述三个前提都得到了实测的肯定或支持. 这进展为猜测性的结构形成理论奠定了很牢靠的基础.

下一步最需要的就是对初始扰动谱的实测确定了. 上面已提到,有两个研究组,一个被称为 BOOMERanG,另一个叫 MAXIMA. 它们从气球上测量了微波背景,测定的是从 10′到 10°范围的扰动. 这测量已对宇宙整体参量的确定起了重要作用,但对提供原初扰动谱讲仍不够. 他们的研究实际上是两个更大的研究计划的预演. 两个卫星观测方案正在准备中. 一个简称为 MAP,约在近年内发射. 另一个叫普朗克巡天器,计划在 2007 年发射. 它们的重要目标(之一)是在背景辐射上细致地测定一切感兴趣尺度上的小扰动. 一旦这目标能达到,那么把结构形成当初值问题处理的

条件就完全成熟了.

总之,宇宙的结构如何形成是宇宙学必须回答,但是也十分难回答的问题.从最近十年的进展看,得到可靠答案的条件已接近具备.我们可望在不远的将来能看到最终的成功.达到这地步,宇宙学也成为一门成熟的学科了.

附录 I　广义相对论提纲①

1．惯性系问题

宏观物理规律追到根本，只有两类：

实物（由原子或分子组成）	服从牛顿动力学
电磁场或电磁辐射	服从麦克斯韦的电动力学

这些规律对什么参考系成立？

在任一惯性系中均成立．

一切惯性系中物理规律都一样　　这叫相对性原理

这基本观念是伽里略在维护日心说时提出的．

牛顿在力学中论证了它．

爱因斯坦在电动力学中论证了它，并由此断言：

所有基本物理规律都将服从相对性原理．

——**在一切惯性系都成立，但在非惯性系不成立．**

这是狭义相对论的主要思想．

此后，人们研究微观物理规律时，（因无经验规律可借鉴）

用了它，取得很大成功．

作为原则性问题，这里有疑点：

惯性系怎么定义？

为什么惯性系会在表达物理规律中有特殊优越的地位？

爱因斯坦意识到这竟然是一个无法回答的问题．

这是他不满意于狭义相对论，

而继续去探索广义相对论的主要动机之一．

---

① 这是为清华大学一年级大学生讲解广义相对论时用过的提纲，现在做了些修改．

快一百年过去了，让我们回头考虑，这问题的要害在哪儿.

这是理解广义相对论所必须弄请的第一个问题.

先看问题的实践方面：

若能从实际的角度指出，哪一个是基本的惯性系.

这就行了.

但实际上只能指出近似的惯性系，找不到严格的惯性系.

地面参考系是一个有用的近似惯性系.

地面惯性系的近似性来自它的转动.

这转动是相对于太阳的.

太阳是严格的惯性系吗？

不！太阳在绕银河系的中心转动.

银河系是惯性系吗？　不！

由于万有引力的存在，**任何参考系**都受力，都有加速度.

严格的惯性系是自然界中不存在的.

再看问题的理论方面：

基本参考系的概念最初来自牛顿的**绝对空间**.

相对性原理使绝对空间成了**玄学**概念，

而物理学是一门实践科学.

原则上不可由观测判定的绝对空间概念在物理上没有意义.

若抛弃绝对空间的概念，惯性系就无法定义了：

——参考系本身没有加速度的是惯性系.（多数人认为）

不行！　有没有加速度以什么为基准？

—$F=ma$　成立的参考系叫惯性系.

不行！　那么动力学规律是什么？

作为原则性的回答，这都不行.　会出现逻辑上的同语反复！

爱因斯坦**不能接受**：

对表述基本物理规律最显优越的竟是一种

理论上无法定义，

实际上不真正存在的参考系！

这被称为“**惯性系疑难**”． 出路何在???

爱因斯坦的想法是

任一种实际的参考系在表述物理规律上都应平等．

——这是他建立广义相对论时的基本想法．

显然的疑问： 能符合事实吗？

从力学讲，**怎么看待非惯性系中要引入惯性力**？

爱因斯坦由此发现了等效原理．

2．等效原理

爱因斯坦是从电磁学问题入手研究相对论的．

犹如：库仑定律不是电磁学一般规律，

运动电荷并不满足它．

他意识到，牛顿引力定律也不是引力的一般规律，

运动物体的引力很可能不满足它．

宏观物理涉及实物和场，场有电磁场，**还有引力场**．

若要**完全奠定宏观物理的理论基础**，

须再找出**引力场**的**一般规律**．

什么是惯性系？

什么是一般的引力定律？

这是爱因斯坦做完狭义相对论后想解决的两个大问题．

广义相对论一箭双雕地解决了．

基本思想体现在**等效原理**之**中**．

爱因斯坦根据一个长久存在而无人解释的事实，产生了一个想法：

惯性力是引力的一种形式．

他注意到的事实是

$$m(\text{惯性})=m(\text{引力}).$$

这涉及到又一个很深刻的物理问题： **什么是质量**．

首先认识到这问题的仍是牛顿.

牛顿动力学中出现质量，描述物体不易被加速的程度，

牛顿万有引力定律中出现质量，描述引力"荷".

没有理由能说明：

不易被加速的物体应当能产生(或受到)更大的引力.

牛顿是物理学家也是思想家.

他知道首先要用实验来澄清这是否是事实.

传说伽里略用斜塔实验证实过这件事，但这仅是传说.

牛顿专门为此做了单摆实验，结果是

$$\frac{m(\text{惯性})}{m(\text{引力})}=1\pm O(10^{-3}).$$

19 世纪末，Eötvös 用扭摆重新作了测定，结果是

$$\frac{m(\text{惯性})}{m(\text{引力})}=1\pm O(10^{-8}).$$

这新实验给当时的爱因斯坦留下了深刻的印象：

m(惯性)$=m$(引力)　**尽管难以理解，但它是一个事实.**

爱因斯坦结合他头脑中的两大问题，产生一个看法：

与 e 是电荷相当，**m(引力)应是牛顿引力荷，**

而　　　　　　　　**m(惯性)本质上是惯性力荷.**

这两种荷的相等说明，惯性力和牛顿引力是不可区分的.

应当把两者看成是同一种力(广义的引力)的不同表现.

这想法就是他的等效原理的雏型.

有先例可借鉴：静电力和磁力是电磁力的不同表现.

作为思考：如果一切微观粒子都有相同的 e/m，

我们能区分静电力和引力吗？

等效原理是广义相对论的思想核心.

爱因斯坦用假想的"封闭舱"实验阐述过他的想法.

若舱内的任何自由落体有共同的加速度(如 $g=8\ \mathrm{m/s^2}$).

怎么推测它的起因？

若 m(惯性)$\neq m$(引力)， 这问题能找到回答.

若能找到一个惯性系作加速度的基准，

这问题也能找到回答.

但这两个前提都不成立，因此我们不可能回答：

它来自引力，还是惯性力？ 还是两者都占一部分？

这是无法判断的.

爱因斯坦由此推断：

惯性力就是引力. 两者不可能区分，也不必区分.

我们将把它叫"广义的引力". (类似广义静电力即电磁力)

问题是须找广义引力的一般规律.(类似于找麦克斯韦方程)

实际上，广义相对论的主要内容正是这一点.

若这问题解决，则惯性系疑难也就解决了.(一箭双雕)

值得强调，能这样看，前提是两条：

m(惯性)$=m$(引力)，

惯性系不存在，参考系的加速度没有明确的基准.

让我们分析，接受等效原理意味着什么.

若只从**力学问题**讲：

(1) 如等效原理对，那么力学规律对任一参考系都一样.

不必分惯性系和非惯性系.

差别仅在两种情况下的引力不同.

有电磁现象作先例，这并不奇怪.

(2) 牛顿引力和惯性力都是一般引力的特殊形式，

问题必须涉及**非力学的物理**现象.(宏观上只有电磁现象)

若借助非力学现象能原则上区分惯性系和非惯性系，

参考系的加速度，以及惯性力就可有确切的定义！

惯性力与引力就能区分. ——等效原理不成立！

参考系的加速度对电磁现象的影响没做过实验.

若等效原理成立，这种影响应归之于引力.

由理论分析（假想实验）知，参考系的加速度的影响肯定存在.

例：惯性系中光沿直线传播.在非惯性系则必沿曲线传播.

这表明，麦克斯韦方程肯定不能对一切参考系都同样地成立.

由此看来， 若接受等效原理，就必须认为：

电磁现象一定也要受引力影响.

人们从经验中总结出来的麦克斯韦方程

只能是没有计及引力影响的电磁规律.

按这思路，爱因斯坦对惯性系获得了与以前很不一样的理解.

什么是惯性系？

引力处处为零的参考系才是真正的惯性系.

这样的惯性系存在吗？

严格讲， 现实世界中是不存在的.

实用讲， 存在.

对电磁现象，引力效应很弱的都是近似惯性系.

回到原则性的高度看问题，

若参考系自身在外引力下作自由下落运动，

它必是一个惯性系.

由于实际引力场的不均匀——惯性系只能是局域的.

且局域惯性系也只能是近似的.

这看法使惯性系疑难得到了满意的解决：

惯性系不具有原则上的特殊性，只有实用上的优越性.

因为它排除了实际上是次要的引力影响，规律简单了.

狭义相对论中的物理规律仅是不计及引力时的规律.

因此如人们常说：**广义相对论本质上是一般性的引力理论.**

它的基本观念已体现在等效原理之中.

为建立广义相对论的理论框架，需要进一步解决两个问题：

（1）**找出广义引力场满足的微分方程.**

（2）**找到引力影响任何非力学现象的规律.**

在广义相对论中，

前一问题的答案是爱因斯坦在 1915 年建立的**场方程**.

后一问题的答案是**广义协变性原理**.

3. 关于时间和空间

人们也常说：相对论是时间和空间的理论.

相对论确使人们对时间和空间的认识发生了重大变革.

但这是副产品.

狭义相对论的要点是抓住了光速的不变性.

Minkowski 首先意识到，它表明：

时间和空间本身都只是影子，它们的联合才是一个整体.

于是人们开始用四维空间来描述这个整体，

这就是 Minkowski 时空.

它依然是一个与物质世界无关的平坦时空.

广义相对论的**要点**是抓住了牛顿引力与惯性力的等效性.

作为**等效原理**的推论，

一般的引力必体现为时空的弯曲.

由于现实世界中引力的不均匀，

现实的四维时空只能是弯曲的.

论证它并不困难，问题在于需要的数学准备太多.

因此我们只讨论它的后果：

(1) 若某时空区内没有(广义)引力，那么它是平坦的.

狭义相对论引用的时空是平坦的，

表明它讨论的是不计引力时的物理学.

又一次看到：**不存在引力的参考系才是惯性系.**

有引力的时空区是弯曲的.

(2) 引力对任何物理现象的影响通过时空的弯曲来实现.

(惯性力只是力学概念，它对非力学现象没有意义)

不同时空点上弯曲程度的不同反映引力强弱的不同.

(3) 引力的强弱应由物质分布及其速度分布决定.

因此， **时空弯曲程度的分布是由物理世界决定的.**

现实时空的几何状况不是数学问题，而是物理问题.

广义相对论把几何学变成了物理学的一部分.

最后看，什么是广义相对论？

人们常说： 广义相对论是一种时空理论.

也常说： 广义相对论是一种引力理论.

这两种说法都对，只是角度不同.

归根到底，广义相对论揭示了

引力场通过时空的结构来体现.

因此，时空与物质是相互紧密地联系着的：

Matter tells spacetime, how to curve.

Spacetime tells matter, how to move.

这就是**广义相对论的基本思想.**

它实际上是**物质和时空的关系**的理论.

附录 Ⅱ　粒子物理大意[①]

1. 什么是粒子物理

人们先是由化学的研究知道，世上纷呈多样的物体都由分子组成. 分子有上百万种而组成分子的化学元素却只有几十种. 这结果暗示人们，若剖析到微观层次，物质世界可能是很简单的. 到 19 世纪末，物理学的研究开始向微观领域发展. 实验证实，元素的原子是由电子绕原子核运动的复合系统. 不同元素的差别在于其原子核有不同的正电荷. 20 世纪 30 年代，中子的发现使人们进一步认识到，几十种元素的上百种原子核（包括同位素）都由质子和中子两种粒子组成. 除了宏观物体之外，人们所知道的物质形态还有光波，无线电波以及 X 射线和 γ 射线等. 当时已清楚，光子是这类物质的共同本原. 这样看来，追溯到微观本原，世界确实变得非常简单. 我们知道的一切物质都可还原成四种粒子，即质子、中子、电子和光子. 人们自然地把这些粒子称为基本粒子，意指它们是构筑世上万物的基本砖块.

可是事情实际上远没有这么简单. 几乎在中子发现的同时，人们立刻意识到，基本粒子世界比上面的描述要复杂. 下面让我们提到早期（30 到 40 年代）的几个重要发现.

正电子的发现. 正电子是在记录宇宙射线的云雾室中被发现的. 它与电子有相同的质量，但却带着相反的电荷. 已知原子中的电子都带负电荷，因此正电子不是宏观物体的组元. 但它的性质却表明正电子与电子同样地基本. 这使当时的人们非常惊讶. 现在粒

① 这是为自学考试教材《物理》（丁俊华主编）所撰写的专题选读材料，这里做了些修改.

子物理学家已彻底弄清，任何基本粒子都有它对应的反粒子存在. 正电子的发现仅是其第一例而已.

中微子的发现. 原子核作 β 衰变时，人们看到的是电子的放出. 但是通过能量和角动量守恒的分析表明，原子核应同时放出一个很轻的中性粒子，它被称为中微子. 这理论的提出在 30 年代，而中微子的实验发现则是 50 年代初的事. 注意在 β 衰变发生之前，原子核中并不存在中微子，因此中微子也不是宏观物体的组元. 它是在衰变的过程中产生出来的. 现在人们知道，基本粒子之间的相互转换是微观世界相互作用的普遍形式.

介子的发现. 在 40 年代，宇宙射线的观测研究中又先后发现了两种质量介乎电子和质子之间的新粒子. 所以被分别称为 μ 介子和 π 介子. 这两种粒子也不是物质的组元. 后来人们在深入的研究后才认识到，这两种粒子的性质很不一样. 较轻的 μ 介子与电子的性质很近，都属于轻子类. π 介子则与质子和中子一样是强子类的一员. 后来人们把介子的名称专用来称呼强子类的成员，因而把 μ 介子改称为 μ 子. μ 子和 π 介子都是不稳定粒子. μ 子的寿命是百万分之几秒. π 介子的寿命更短，仅有万万分之几秒. 粒子实验已经证实，不稳定性是绝大多数基本粒子的共性. 稳定的基本粒子只有极少几种.

从上面讨论到的几个例子看，所发现的新粒子都并不是物体的组元粒子，但却都与组元粒子同样地基本. 这些发现使人们逐渐认识到，基本粒子世界仍然很复杂. 作为宏观物体组元的四种粒子仅是其中的一小部分而已. 于是从 20 世纪中叶起，物理研究又多了一个新的分支. 这就是基本粒子物理.

在发明了加速器之后，新粒子的发现效率大大地提高了. 到 60 年代初，实验发现的基本粒子的数目已过百种. 而且显然，随着加速器能量的提高，还会有大量的新粒子会被发现出来. 原来人们期望基本粒子的研究会给物质世界描绘出一幅很简明的图象. 岂知结果却相反，基本粒子的种类比化学元素的种类还多！这使人们

悟到，这些粒子只是物质结构中比原子核更深的一个层次，而并不是物质世界的终极本原. 这样看来，基本粒子不是一个合适的名称. 于是人们去掉“基本”二字，而把它们简称为粒子. 相应的研究领域也改称为粒子物理.

2. 四种相互作用和粒子的分类

物体间的相互作用是造成千变万化的物理现象的根本原因. 因此除物体本身外，物体间的相互作用规律也是物理学的主要研究对象.

在经典物理中，相互作用表现为宏观的力. 力的效果是引起宏观物体的变形或运动的加速等. 微观物体与宏观物体有质的不同，因此相互作用的表现方式也有质的区别. 微观粒子间的相互作用表现为粒子间的转换. 例如中子的 β 衰变是中子转换成了质子，同时产生出电子和(反)中微子. 用写化学反应的方式来写，这过程是

$$n \to p + e^- + \bar{\nu},$$

其中 $\bar{\nu}$ 是反中微子. 作用力的强弱则体现为该转化过程发生的几率的大小.

在经典物理中力的表现形式虽很多样，但在本质上却只有两种，即引力和电磁力. 地面上物体感受到的引力主要来自地球. 它就是通常讲的重力. 此外种类繁多的力，如物体间的摩擦力或弹性力等，本质上都是其分子间的电磁力的表现. 延伸到微观过程，由于粒子的质量太小，引力过于微弱而不产生可观测的效果. 电磁力在带电的微观粒子间则依然在起作用. 但是粒子间除电磁力之外，事实表明一定还有其他的力存在.

原子核是若干个质子和中子的结合体. 质子与质子间有静电斥力，它们为什么能结合成如此紧密的原子核?这明显说明还有一种吸引性的力在起作用. 这种作用必比静电力更强，因而它被称为强作用力. 此外，上面已提到的 β 衰变中涉及不带电粒子，因此它也不是电磁力的效果. 定量分析表明这种作用力比电磁力弱很多. 因

此被称为弱作用力．它又是一种新的作用力．重要的是从各种粒子过程的分析发现，粒子间的作用力除电磁力之外就只有这强力和弱力两种了．这样人们对自然界的相互作用建立了一个概括的认识：微观过程中只有四种相互作用．按强弱排序，它们是强作用力、电磁力、弱作用力和引力．由于强力和弱力只在微观距离上起作用，所以在宏观现象上只有电磁力和引力．

如同不带电的粒子感受不到电磁力一样，实验结果的分析表明，有些粒子完全不受强作用力的影响．上面已提到的电子、中微子和 μ 子都是．这些粒子被统称为轻子类，因为这些代表性粒子都比质子要轻．轻子都不参与强作用，既不产生强力也不感受强力．带电的轻子参与弱作用和电磁作用．不带电的轻子（中微子）则只参与弱作用．此外有一大类粒子以质子、中子和 μ 介子为代表．它们既参与弱作用和电磁作用（若带电），且也参与强作用．由于在几种力都存在时，强作用占压倒优势，因而这些粒子被称为强子类．这样，人们按相互作用的不同，把粒子分成了轻子和强子两大类．此外还有一类叫规范粒子．简单地讲，它们是在粒子间传递作用力的媒介．我们都知道光子是电磁力的传递者．而 $W^{\pm}$ 和 Z^0 则是弱力的传递者．按至今很成功的强作用理论，强力的传递者是胶子．由于胶子是不能单独出现的粒子，因此无法直接观测到它．理论上把引力的传递者叫引力子．它的存在尚没有实测证据．下面列表给出了已观测到的规范粒子和轻子，以及一些代表性的强子的性质．这里把强子又分成了两个亚类——介子和重子．它们的区别将在后面谈到．

经验使人们对粒子及其相互作用作这样的分类．它的价值不仅是使人们对粒子世界的全局有了更清晰的了解，而且对以后的深入研究是重要的导向．

3. 强子的夸克结构

到现在，轻子族的成员已发现了六个．它们是电子、μ子、τ子和

分别与之对应的三种中微子ν_e，ν_μ和ν_τ. 按已为实验广泛证实的粒子物理标准模型，轻子类也只有这些成员了. 轻子类的情况较简单，而强子类的情况却显得很复杂. 现在已发现的强子有八百多种. 它的种数过多是一种迹象，暗示着强子内部还有更深层次的结构.

在化学发展到一定阶段，元素的数目越来越多，门捷列夫总结出了元素的周期性. 它对 20 世纪的人们认识原子的结构起了极关键的导向作用. 如所周知，泡利的不相容原理就是为解释元素周期性而提出的理论. 它是今天认识不同原子的差别的核心. 强子的研究过程有些类似. 在强子种类已很多时，人们开始按唯象性质对不同强子作分类. 60 年代中期，在分类规律的指引下有人提出，强子是若干个夸克组成的复合体，夸克则是更深层次上的组元粒子. 这使我们对微观世界的认识又前进了一步.

按标准模型，夸克共有六种. 粒子物理学家风趣地称夸克有六种不同的"味". 最轻的两种被称为 u 夸克和 d 夸克. u 夸克带的是正电荷，电量是电子电量的 2/3. d 夸克带负电荷，电量是电子电量的 1/3. 最轻的重子仅由 u，d 夸克组成. 质子由 u，u，d 三个夸克组成，其电荷与电子等量反号. 中子由 u，d，d 三个夸克组成，其电荷为零. 其他重子也都由三个夸克组成. 介子则是由一个夸克和一个反夸克构成的复合体. 由 u 和反 d 夸克组成的是带一个正电荷的 π^+ 介子. 由反 u 和 d 夸克组成的是 π^- 介子.

以电子电量为单位，所有已发现的粒子的电量都是整数，而夸克所带的电量却是分数. 这是一个很显著的特征. 于是人们开始以分数电荷为标志来寻找夸克. 几年的努力没有成果，而相应的夸克强相互作用理论却发展了起来. 这就是量子色动力学理论. 按这理论，夸克不仅有六种"味"的区别，而且还有三种"色"的区别. 它们被称为红夸克、蓝夸克和黄夸克. 注意这仅是形象化的叫法. 不同颜色的夸克靠胶子粘合在一起，能造成白色的夸克复合体. 这才是能观测到的强子. 而非白色的单个夸克、胶子或它们的复合体是不能单独出现的. 这样，单个夸克不被发现就成必然的后果了.

初一听，这理论有点像为找不到夸克而编造的遁词. 其实不然. 一个新理论的正确与否，不取决于它解释了什么，而在于它对尚不清楚的事实能作出预言，并且这些预言能为将来的实验所证实. 经过 20 年的实验考验，这个夸克强作用理论已被大量实验证明是正确的. 所以夸克虽没有找到，但现在已很少有人再为此而怀疑它的真实性了.

夸克理论的确立使我们对粒子世界的认识进了一大步. 原来作为原子核组元的质子和中子等强子都不是基本粒子，而是更深层次粒子的复合物. 按今天的认识，夸克和轻子一起，才共同代表一种深层次的组元粒子. 轻子没有强作用，所以它们不能形成结构紧密的复合粒子. 今天的实验现象没有证据显示夸克或轻子还有内部结构. 至于它们是否是物质的终极本原，这是将来的物理学家才能回答的问题. 无论如何，现在人们已不愿再轻言"基本"二字了.

4. 相互作用的统一

前面已提到：微观粒子间的相互作用有四种，它们各自是独立的. 其中有两种是短程作用，对宏观现象起不了作用，因此宏观力的本原只有引力和电磁力. 理论家常会有理想化的倾向，他们常想在表面上很不同的事情背后找出简单的统一本原. 爱因斯坦在建立了相对论之后，曾致力于研究引力和电磁力的统一. 他想给这两种表观上没有联系的作用力作出统一的描述. 虽然当时爱因斯坦的名望如日中天，可是他为此花费了很大的精力，最后在这领域中却一无所获. 物理学是一门实验科学. 一个理论是否正确的惟一标志是它是否符合事实. 因此理论家追求统一的念头能否实现，取决于他们的这种想法是否与客观情况相一致. 而这点是很难事先判断的.

在后来的理论家看来，爱因斯坦的失败不是由于他追求统一的想法不对，而是由于他错误地想在宏观物理的基础上寻求统一.

宏观物理规律是唯象性的，而不是本原性的. 因此在微观粒子的动力学理论确立后，追求几种不同的相互作用相统一的努力又复活了起来. 在 60 年代后期，终于有人成功地迈出了第一步. 这就是弱力和电磁力统一理论的建立. 它当然是人类认识微观世界上的又一重大成果.

从现象上看来，弱力和电磁力是性质差别很大的两种力. 前者是短程力而后者是长程力. 它们在强度上的差别有十个量级之大. 这样两种力怎么会在本质上是同一种?由量子场理论看，弱力的微弱性和它的短程性来自同一个原因，那就是传递弱力的媒介粒子很重. 现在已成功的弱电统一理论的要点是指出，在能量很高(远大于几百吉电子伏)的现象中，传递弱力的媒介粒子与光子一样是质量为零的. 作为某种内部对称性的后果，弱力和电磁力原是同一种力. 现在人们就简单地把这种统一的力叫作弱电力. 在能量较低时，一种被称为希克斯机制的物理效应把该种内部对称性破坏了. 相应的效果之一是使某些传递力的媒介粒子获得了很重的质量. 它们就是表 1 中列出的 $W^{\pm}$ 和 Z^0 粒子. 这些粒子传递的力变得很微弱，力程变得很短. 这就是我们在低能现象上看到的弱力. 有一种媒介粒子依然保持质量为零. 它就是传递电磁力的光子. 这样，弱力和电磁力才在表观上显出了很大的不一样. 这理论在术语上称为有对称性自发破缺的规范场理论. 传递力的媒介粒子因此被叫做规范粒子.

当然，理论的设计总有一定的人为性. 但如我们已强调，这理论能否被大家接受为正确的理论，取决于它的一系列推论能否得到事实的肯定. 弱电统一理论从 70 年代起已经受住了许多关键性的实验考验，从而在 80 年代初获得了诺贝尔物理奖. 后者是一个标志，说明它得到了人们的首肯.

弱电统一理论的成功大大地鼓舞了有理想化倾向的理论家们. 他们开始想把强作用力与弱电力再进一步统一起来，甚至想把引力也统一在内. 如果这种努力成功了，那意味着自然界的基本相互作用本质上只有一种. 这确实应当被认为是一件了不起的事. 但

是检验这种理论很难,因此人们至今还不知道这种理论是否对.看来,理论家想极终地探索物质及其相互作用本原的努力很可能是会成功的.无论如何,他们前面还有一段漫长的路要走.

表 1　粒子分类和性质简表

分类		粒子名称	电荷（以电子电量为单位）	自旋（以$\hbar$为单位）	质量（mc^2以MeV为单位）	寿命（以s为单位）	主要衰变方式
规范粒子		γ(光子)	0	1	0	∞	
		W	±1	1	81800	$\sim10^{-25}$	$W^+\to e^++\nu_e$
		Z	0	1	92600	$\sim10^{-25}$	$Z^0\to e^++e^-$或$\mu^++\mu^-$
轻子		ν_e(e中微子)	0	1/2	$<4.6\times10^{-6}$	∞	
		ν_μ(μ中微子)	0	1/2	<0.25	∞	
		ν_τ(τ中微子)	0	1/2	<27.7	∞	
		e(电子)	−1	1/2	0.511	∞	
		μ	−1	1/2	106	2.20×10^{-6}	$\mu^-\to e^-+\nu_e+\nu_\mu$
		τ	−1	1/2	1784	3.3×10^{-13}	$\tau^-\to\mu^-+\bar{\nu}_\mu+\nu_\tau$
强子	介子	π	±1	0	140	2.60×10^{-8}	$\pi^+\to\mu^++\nu_\mu$
			0	0	135	0.87×10^{-16}	$\pi^0\to\gamma+\gamma$
		K	±1	0	494	1.24×10^{-8}	$K^+\to\mu^++\nu_\mu$
			0	0	498	K_S^0 0.89×10^{-10} K_L^0 5.18×10^{-8}	$K_S^0\to2\pi$ $K_L^0\to3\pi$
	重子	p(质子)	+1	1/2	938	∞	
		n(中子)	0	1/2	940	898	$n\to p+e^-+\bar{\nu}_e$
		Λ	0	1/2	1116	2.63×10^{-10}	$\Lambda\to p+\pi^-$或$n+\pi^0$

附录Ⅲ　自然单位制

物理学中出现了三个很基本的普适常数，它们是光速 c，普朗克常数 $\hbar$ 和玻尔兹曼常数 k. 光速其实是一切零质量粒子的速度，也是自然界的速度上限. 普朗克常数是一切量子规律中的基本单元. 玻尔兹曼常数是热现象中宏观量与微观量间的桥梁. 讲这些量很基本，指在相关的公式中它们一定要出现的. 可是它们的值都很复杂. 这数值上的复杂是我们随便地规定了时间、长度和质量的单位所造成的. 如果适当地选择单位，让这些常数的值都为 1，显然是能办到的. 这就是建立自然单位制的想法. 做到这点后，相关公式中这三个常数都自然不再出现. 它不仅使公式简单，而且使物理量间的关系也更明白.

一个完整的单位制要包含四个基本量. 普通单位制中取长度、时间、质量和温度. 在宇宙学的自然单位制中除把光速 c、普朗克常数 $\hbar$、玻尔兹曼常数 k 都作为基本量，并规定为 1 外，再把能量单位规定为 1 GeV. 这些量在普通单位制中的数值是

$$1\ c = 2.9979\times10^{8}\ \mathrm{m\cdot s^{-1}},$$

$$1\ \hbar = 1.0546\times10^{-34}\ \mathrm{J\cdot s},$$

$$1\ k = 1.3807\times10^{-23}\ \mathrm{J\cdot K^{-1}},$$

$$1\ \mathrm{GeV} = 1.6022\times10^{-10}\ \mathrm{J}.$$

为了方便，这里用了能量的普通单位焦耳，记作 J.

记得 $1\ \mathrm{J} = 1\ \mathrm{kg\cdot m^{2}\cdot s^{-2}}$.

很容易用这四个关系把时间、长度、质量和温度反解出来. 因取 $c=\hbar=k=1$，所以任何量都以 GeV 为量纲：

时间　$1\ \mathrm{s} = 1.5192\times10^{24}\ \mathrm{GeV^{-1}}$.

长度　$1\ \mathrm{m} = 5.0677\times10^{15}\ \mathrm{GeV^{-1}}$.

质量　$1\ \mathrm{kg}=5.6095\times10^{26}\ \mathrm{GeV}$.

温度　$1\ \mathrm{K}=8.6170\times10^{-14}\ \mathrm{GeV}$.

这就是四个基本量的普通单位与自然单位的关系.若需要反过来从这些量在自然单位下的值转化为普通单位下的值,则有

时间　$1\ \mathrm{GeV}^{-1}=6.5822\times10^{-25}\ \mathrm{s}$.

长度　$1\ \mathrm{GeV}^{-1}=1.9733\times10^{-16}\ \mathrm{m}$.

质量　$1\ \mathrm{GeV}=1.7827\times10^{-27}\ \mathrm{kg}$.

温度　$1\ \mathrm{GeV}=1.1605\times10^{13}\ \mathrm{K}$.

这样,两种单位制间的转换关系已完全清楚了.

对于其他物理量,只要记得它在普通单位制中的量纲,那么在自然单位制中的量纲就立即有了.让我们记得几个宇宙学中的常用量的自然量纲:

时间、长度	量纲为 GeV^{-1}.
质量、能量、温度	量纲为 GeV.
体积	量纲为 GeV^{-3}.
粒子数密度	量纲为 GeV^{3}.
质量密度、能量密度、压强	量纲为 GeV^{4}.
牛顿引力常数 G	量纲为 GeV^{-2}.

要在数值结果上相互折算也很容易.例如

粒子数密度	$1\ \mathrm{GeV}^3=1.3014\times10^{47}\ \mathrm{m}^{-3}$.
质量密度	$1\ \mathrm{GeV}^4=2.3201\times10^{20}\ \mathrm{kg\cdot m^{-3}}$.
能量密度	$1\ \mathrm{GeV}^4=1.3014\times10^{47}\ \mathrm{GeV\cdot m^{-3}}$.
普朗克能量	$m_{\mathrm{pl}}=G^{-1/2}=1.2211\times10^{19}\ \mathrm{GeV}$.
普朗克时间	$t_{\mathrm{pl}}=G^{1/2}=8.1893\times10^{-20}\ \mathrm{GeV}^{-1}$ $=5.3904\times10^{-44}\ \mathrm{s}$.

宇宙学问题涉及相对论、量子物理和热学,有关公式中必有 c, $\hbar$ 和 k 的若干次幂.它们实际上没有告诉我们任何物理内容,而只是累赘.所以人们常采用自然单位制讨论,以避免出现这种累赘.以下就我们用到的方面作些讨论和说明.

相对论认识到时间和空间是一个统一的整体. 取光速 c 为 1 的首要好处是使时间和长度有同样的量纲. 这样做后出现了例如公式(4.39): 视界大小 $l_h(t)=2t$. 它应理解为这时间乘光速就是视界大小,而光速大小是 1.

取光速为 1 的另一好处是把质能关系简化成了 $E=m$. 在概念上它直接告诉我们,描述质量和能量的是同一个物理量,两者不必区分. 实用上因此也很方便. 粒子物理正因这道理,用 GeV 来描述粒子的(静)质量. 我们也同样沿用. 按同样的道理,(质量)密度和能量密度也是同一个物理量. 宇宙学需要把辐射和实物统一讨论,这样做是方便的.

在广义相对论中,能流密度、动量密度和动量流密度也都是引力的源. 现在它们与通常意义下的密度有同样的量纲,很便于统一处理. 压强作为动量流密度,与密度有一样的量纲. 由此才会在公式(2.11)中出现密度与压强的直接相加. 对辐射物质,压强与密度的关系简化成了 $P=\rho/3$. 在普通单位下还应有一个因子 $1/c^2$.

温度的基本意义是描写热运动的剧烈程度. 按普通单位制,在温度为 T 的气体中,粒子的热动能为 kT 的量级. 取玻尔兹曼常数 k 为 1,则温度 T 直接描写了粒子的热动能. 这其实更符合温度概念的本意. 在这样做后,温度与能量、质量都有了同样的量纲,对于需要比较时是方便的. 例如在早期宇宙中,质量为 m 的粒子会热碰撞产生的条件是 $T\gg m$. 用普通单位下的话讲,应是 $kT\gg mc^2$. 宇宙学中用电子伏 eV 描写温度,利用的就是这好处. 需要记得的是: 温度 1 eV$\approx 10^4$ K.

把普朗克常数$\hbar$取作 1,给粒子理论公式带来很多简便. 这我们不讨论. 我们主要是在导出辐射气体的密度(见(4.9)式)和数密度(见(4.15)式)时用到了它,因为辐射物质是量子气体. 其后果是使密度(或数密度)与温度的关系简单了. 例如在温度为 1 MeV 时近似地有 $\rho=T^4$. 用 GeV 为基本量纲,立即得到 $\rho=10^{-12}$ GeV^4. 用1 GeV^4=2.3201×10^{20} kg/m^{3}折合成普通单位,它是10^8 kg/m^3.

其物理含义就大体清楚了. 如果用普通单位写，同样作略去一个量级为 1 的因子后的公式是 $\rho=T^4\hbar^{-3}k^4c^{-3}$. 这里的物理要点是 ρ 与 T^4 正比，这不会受单位制影响. 只是普通单位制中出现了很复杂的比例系数. 对比之下，我们能看出用自然单位制写公式的好处. 但是两者的差别也仅在这里.

附录Ⅳ 天文学和宇宙学常量

天文学量

太阳质量 $M_{\odot}$	1.989×10^{30} kg
太阳半径 R	6.960×10^{8} m
太阳光度 L	3.90×10^{26} J·s^{-1}
恒星质量范围	0.1—$10^{2}\ M_{\odot}$
星系质量范围	10^{6}—$10^{13}\ M_{\odot}$
星系大小	10^{0}—10^{2} kpc
星系团平均大小	5 Mpc
星系团内平均星系数	10^{2}
超团和大空洞大小	50 Mpc

宇宙学量

哈勃常数 H_0	$100\ h$ km·s^{-1}·Mpc^{-1}
	$h=0.5$—0.8
哈勃时间 H_0^{-1}	$9.78\times10^{9}\ h^{-1}$ a（即 yr）
哈勃距离 cH_0^{-1}	$3.00\times10^{3}\ h^{-1}$ Mpc
临界密度 ρ_c	$1.88\times10^{-26}\ h^{2}$ kg·m^{-3}
背景光子温度 $T_{\gamma 0}$	2.728 ± 0.004 K
背景光子数密度 $n_{\gamma 0}$	412 cm^{-3}
背景光子密度 $\rho_{\gamma 0}$	4.67×10^{-31} kg·m^{-3}
背景光子能量密度 $\rho_{\gamma 0}$	2.62×10^{-7} eV·m^{-3}
光子退耦时刻 $t_{\gamma dec}$	$1.79\times10^{5}(\Omega_0 h^{2})^{-1/2}$ a
光子退耦温度 $T_{\gamma dec}$	0.258 eV
背景中微子数密度 $n_{\nu 0}$	112 cm^{-3}

背景中微子温度 $T_{\nu 0}$	1.95 K
实物与辐射等量时刻 t_{eq}	$1.32\times10^{3}(\Omega_0 h^2)^{-2}$ a
实物与辐射等量温度 T_{eq}	5.64 $(\Omega_0 h^2)$ eV